University of California

Los Angeles

The Geometry/Gauge Theory Duality and the Dijkgraaf–Vafa Conjecture

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Physics

by

Masaki Shigemori

2004

The dissertation of Masaki Shigemori is approved.

University of California, Los Angeles

2004

To my parents

Table of Contents

List of Figures

List of Tables

Acknowledgments

First of all, I would like to express my gratitude to my thesis advisor Per Kraus, for sharing with me his insights and intuition, and for teaching me attitude toward physics — how to understand and think about things physically. I also wish to thank my other advisor Eric D'Hoker, who always explained to me any physical concepts in his usual crystal-clear manner. It is also my pleasure to thank Changhyun Ahn, Bo Feng, Yutaka Ookouchi and Anton Ryzhov for exciting collaborations, and Iosif Bena, Brian Forbes, Michael Gutperle and Sebastian de Haro for enlightening discussions.

I am grateful to my parents, my sister, and my wife for their unfailing support and for giving me the opportunity to carry out my Ph.D. studies.

Chapter 2 is a version of P. Kraus and M. Shigemori, "On the matter of the Dijkgraaf–Vafa conjecture," JHEP **0304**, 052 (2003) [arXiv:hep-th/0303104] [23]. We would like to thank Iosif Bena, Eric D'Hoker, and Anton Ryzhov for helpful discussions.

Chapter 3 is a version of P. Kraus, A. V. Ryzhov and M. Shigemori, "Loop equations, matrix models, and $\mathcal{N} = 1$ supersymmetric gauge theories," JHEP **0305**, 059 (2003) [arXiv:hep-th/0304138] [31].

Chapter 4 is a version of K. Intriligator, P. Kraus, A. V. Ryzhov, M. Shigemori and C. Vafa, "On low rank classical groups in string theory, gauge theory and matrix models," Nucl. Phys. B **682**, 45 (2004) [arXiv:hep-th/0311181] [87]. I would like to thank F. Cachazo for valuable discussions.

Chapter 5 is a version of C. Ahn, B. Feng, Y. Ookouchi, and M. Shigemori, "Supersymmetric Gauge Theories with Flavors and Matrix Models," Nucl. Phys.

B **698**, 3 (2004) [arXiv:hep-th/0405101] [98]. I would like to thank Freddy Cachazo, Eric D'Hoker, Ken Intriligator, Romuald Janik, Per Kraus, and Hirosi Ooguri for enlightening discussions.

This research was partly supported by Heiwa Nakajima Foundation Scholarship and NSF grant PHY-0099590.

Abstract of the Dissertation

The Geometry/Gauge Theory Duality and the Dijkgraaf–Vafa Conjecture

by

Masaki Shigemori

Doctor of Philosophy in Physics

University of California, Los Angeles, 2004

In this dissertation we discuss various issues concerning application of the Dijkgraaf–Vafa conjecture to the study of supersymmetric gauge theories. The conjecture states that for a large class of $\mathcal{N}=1$ supersymmetric gauge theories, the exact effective superpotential for the glueball superfield can be computed by matrix models. This approach is very powerful in that it provides a systematic way of computing the nonperturbative, sometimes even exact, superpotential of the system, which was possible only on a case-by-case basis in the more traditional approach based on holomorphy and symmetry.

This conjecture has been checked for many nontrivial examples, but the range of applicability of the conjecture remained unclear. In Chapter 2, we give an explicit example, $Sp(N)$ theory with antisymmetric tensor, which challenges the applicability of the conjecture. We will show that, the superpotential obtained by the Dijkgraaf–Vafa approach starts to disagree with the standard gauge theory result at $N/2+1$ loops. Thus we present a relatively simple example for which a straightforward application of the Dijkgraaf–Vafa conjecture leads to a different result from the known result. Furthermore, in Chapter 3, we will reproduce

the same discrepancy in the generalized Konishi anomaly method, an alternative approach to computing the glueball superpotential.

In order to look for the physical origin of the discrepancy, in Chapter 4, we consider the string theory realization of the gauge theories by certain Calabi–Yau compactifications, on which the Dijkgraaf–Vafa conjecture is based. By closely analyzing the geometric transition of the Calabi–Yau space, we will uncover the following prescription regarding when to include a glueball field: one should not include a glueball field for $U(0)$, $SO(0)$, $SO(2)$ while for all other gauge groups, including $U(1)$ and $Sp(0)$, one should consider an associated glueball field. We will explicitly show that the discrepancy found in Chapters 2 and 3 is resolved if we follow this prescription and introduce a glueball field for the "$Sp(0)$" group.

In Chapter 5, we consider generalizing the result in Chapter 4 to include flavors; we give a prescription regarding when to include glueball fields based on string theory realization, and demonstrate that the matrix model computations along with the generalized prescription correctly reproduce the gauge theory results.

CHAPTER 1

Introduction

1.1 Supersymmetric gauge theory

The importance of the study of supersymmetric gauge theory cannot be emphasized enough. Supersymmetric gauge theory is widely regarded as the most promising and natural candidate for a framework which extends the Standard Model and which may be able to explain the outstanding issues of the Standard Model such as the hierarchy problem, grand unification, proton decay, and possibly the cosmological constant problem [1]. Supersymmetric gauge theory is attractive not only from such phenomenological viewpoints but also from theoretical viewpoints. In particular, supersymmetry makes the theory more tractable than the theory without supersymmetry. For example, in the ordinary, non-supersymmetric QCD, we do not have an analytic way to study even the most fundamental properties of the theory such as confinement or chiral symmetry breaking, because of its notorious strongly coupled dynamics. However, in its supersymmetric versions, those properties can be proven exactly. The reason why we have analytic control over supersymmetric theory is closely related to the holomorphy of the superpotential which supersymmetric theories possess.[1] Therefore, supersymmetric gauge theory is important in that it serves as a toy model which helps us to understand the properties of non-supersymmetric theory

[1]For reviews on supersymmetric gauge theories, see e.g. [2, 3, 4].

in a simpler setting.

Many powerful techniques have been developed for analytically studying the properties of supersymmetric gauge theories. Among the physical quantities of the theory, superpotential is of fundamental importance because it determines the structure of the vacua of the theory. The superpotential can sometimes be computed exactly using holomorphy and symmetry considerations, however there was no systematic way of computing it — it has been worked out only on a case-by-case basis.

1.2 The Dijkgraaf–Vafa conjecture

Recently Dijkgraaf and Vafa [5, 6, 7] proposed a powerful, systematic way of computing superpotential — the so-called Dijkgraaf–Vafa conjecture. For a large class of $\mathcal{N} = 1$ supersymmetric gauge theories, this conjecture states: i) at low energies, the holomorphic physics, namely the superpotential term in the action, is captured by the glueball superfield S, ii) the exact effective superpotential of the glueball S can be computed by $0 + 0$ dimensional field theory, i.e., matrix model. As we will review later (section 4.3), this conjecture was proposed based on the geometric transition (conifold transition) duality [8, 9, 10] in the string theory realization of the models by Calabi–Yau compactification, and on the fact that the four-dimensional superpotential of such a compactified theory can be computed by topological strings [11].

The first part of the conjecture has a clear physical meaning, although it is hard to prove: at low energies, the strongly coupled gauge theory confines and the vacuum acquires nonzero expectation value of the glueball field, $S = -(1/32\pi^2)\mathrm{Tr}[\mathcal{W}_\alpha \mathcal{W}^\alpha]$. Here $\mathcal{W}_\alpha$ is the field strength superfield, and the low-

est component of S is the fermion (gluino) bilinear whose nonvanishing vacuum expectation value (vev) signals chiral symmetry breaking. The low energy excitation around this vacuum should be the glueball field S changing slowly over spacetime.

The second part of the conjecture claims that the momentum integration in the computation of Feynman diagrams in supersymmetric gauge theory is in fact trivial, and reduces simply to computation of combinatorial factors. Furthermore, almost all such diagrams actually vanish — for example, for $U(N)$ theory with a chiral superfield in the adjoint representation, only diagrams whose topology in the 't Hooft double line notation [12] is S^2 are nonvanishing. Later, this second part of the conjecture was proven [13, 14, 15] totally in the standard framework of the super-space Feynman diagram expansion [16]. It is surprising that this tremendous simplification had been overlooked until the work of Dijkgraaf and Vafa.

For example, for $U(N)$ theory with an adjoint chiral superfield Φ, this conjecture states that the exact effective glueball superpotential is given by

$$W_{\mathrm{eff}}(S) = N\frac{\partial \mathcal{F}_{S^2}}{\partial S}, \tag{1.2.1}$$

where $\mathcal{F}_{S^2}$ is the contribution to the matrix model free energy from diagrams of S^2 topology.

This conjecture was checked for many nontrivial examples (see [19] for a list of references) including the cases with fundamentals, baryonic interaction, multi-trace interaction; in some theories one can obtain the exact superpotential by more traditional techniques based on holomorphy and symmetry considerations, and compare the results with the ones predicted by the Dijkgraaf–Vafa conjecture. The agreement was perfect.

1.3 Generalized Konishi anomaly approach

Inspired by the work of Dijkgraaf and Vafa, an alternative approach for obtaining the glueball superpotential was developed in [17, 18], based on the generalized Konishi anomaly[2].

In the matrix model approach, one can compute the glueball superpotential from the expectation values of matrices of the form $\langle \mathrm{Tr}[\mathbf{\Phi}^n] \rangle$, where $\mathbf{\Phi}$ is an $\mathbf{N}\times\mathbf{N}$ matrix which corresponds to the adjoint chiral superfield Φ in gauge theory. It is well known that these matrix model expectation values satisfy certain relations called the loop equations, which are analogues of the Schwinger–Dyson equations in field theory. These loop equations are so powerful that they determine those expectation values $\langle \mathrm{Tr}[\mathbf{\Phi}^n] \rangle$ completely, up to a finite number of undetermined parameters.

Because matrix model and gauge theory are closely related according to the Dijkgraaf–Vafa conjecture, there should be some relations also in the gauge theory framework, which are powerful enough to determine the glueball superpotential. In gauge theory, the glueball superpotential can be determined by the vev of the chiral operators, which are defined to be operators annihilated by supersymmetry generators. For example, $\mathrm{Tr}[\Phi^n]$, $\mathrm{Tr}[\mathcal{W}_\alpha \mathcal{W}^\alpha \Phi^n]$ are chiral operators. Chiral operators form a ring called the chiral ring, and their vev are independent of spacetime coordinates. Furthermore, the vev of the product of chiral operators is equal to the product of the vev of the chiral operators. Therefore, chiral operators are just as ordinary numbers, as far as their expectation values are concerned. Obviously, the vev of these chiral operators are natural candidate of the quantities that play the role of $\langle \mathrm{Tr}[\mathbf{\Phi}^n] \rangle$ in matrix model. Indeed, in [17, 18], the Schwinger–Dyson

[2]For a recent review on the generalized Konishi anomaly approach and the diagrammatic approach to the Dijkgraaf–Vafa conjecture, see [19].

equations for these operators were derived and it was shown that these equations determine all vev's of the chiral operators up to a finite number of undetermined parameters. These Schwinger–Dyson equations are called the generalized Konishi anomaly equations, because these generalize the Konishi anomaly equation [20] which is the superfield version of the ordinary anomaly equations.

The generalized Konishi anomaly approach established a firm connection between the standard gauge theory formalism and the Dijkgraaf–Vafa conjecture which appeared to be rather out of a hat from the traditional gauge theory point of view, although it of course is a very natural conjecture from the viewpoint of string theory.

1.4 $Sp(N)$ theory with antisymmetric tensor

As mentioned above, the Dijkgraaf–Vafa conjecture passed many nontrivial checks. However, it is not clear how large a class of supersymmetric theories the conjecture is applicable to; it is very important to investigate the applicability of the conjecture.

$Sp(N)$ theory[3] with a chiral superfield in the antisymmetric tensor is among those theories for which the exact superpotential can be obtained by traditional techniques [21, 22]. As equation (1.2.1) exemplifies, the Dijkgraaf–Vafa conjecture generally predicts a superpotential with a simple pattern in N, the rank of the gauge group. However, in this $Sp(N)$ theory, the dynamical superpotential obtained by the traditional techniques does not appear to have any obvious pattern in N. Therefore, this theory is an ideal touchstone for checking the applicability of the conjecture.

[3] Our convention for $Sp(N)$ is such that N is an even integer, so that $Sp(N) \subset U(N)$ and $Sp(2) \cong SU(2)$.

In Chapter 2, we apply the Dijkgraaf–Vafa conjecture to this $Sp(N)$ theory with antisymmetric tensor based on the diagrammatic approach [13, 14, 15], and the effective glueball superpotential is computed using $Sp(N)$ matrix model, for the cubic tree level superpotential and trivial breaking pattern $Sp(N) \to Sp(N)$. Comparing the resulting superpotential with the one computed using the traditional techniques for $Sp(4)$, $Sp(6)$ and $Sp(8)$, we find that the two superpotentials agree up to $N/2$-th order in the matrix model perturbation theory, with discrepancy setting in at the next order. Thus, this $Sp(N)$ theory gave the first explicit example for which a straightforward application of the Dijkgraaf–Vafa conjecture leads to a different result from the known result. We will discuss possible origin of the discrepancy, e.g., possible ambiguity in defining high powers of the glueball superfield S^n with $n \geq h = N/2 + 1$.

In Chapter 3, we apply the method of the generalized Konishi anomaly to theories based on the classical gauge groups with various two-index tensors and fundamentals, including $Sp(N)$ theory with antisymmetric tensor. We will see that the generalized Konishi anomaly approach gives the same discrepancy as was observed in the matrix model approach. Because the generalized Konishi anomaly approach is not based on diagram expansion, we can discuss the aforementioned S^h problem from a different point of view. Moreover, we present a general formula which expresses the solutions to the generalized Konishi anomaly equation in terms of the solutions to the loop equations of the corresponding matrix model.

1.5 String theory prescription

The discovery of the above "counterexample" to the Dijkgraaf–Vafa conjecture by [23] stimulated active research [15, 24, 25, 26, 27, 28, 29] to look for the physical origin of the discrepancy. In particular, Cachazo [24] showed that the

generalized Konishi anomaly equations governing the $Sp(N)$ theory with antisymmetric tensor matter [30, 31] can be mapped to those of $U(N+2K)$ gauge theory with adjoint matter, with the breaking pattern $Sp(N) \rightarrow \prod_{i=1}^{K} Sp(N_i)$ mapped to $U(N+2K) \rightarrow \prod_{i=1}^{K} U(N_i+2)$. Here, K is the number of the critical points in the tree level superpotential. Especially, the trivial breaking pattern studied [23] for the cubic superpotential, which can be written as $Sp(N) \rightarrow Sp(N) \times Sp(0)$, is mapped to $U(N+4) \rightarrow U(N+2) \times U(2)$. It was shown [24] that this $U(N+2) \times U(2)$ theory correctly reproduces the superpotential of the $Sp(N)$ theory and resolves the discrepancy found in [23]. Because $U(2)$ has a glueball dynamics in it, this map implies that we should consider a glueball field for the "$Sp(0)$". In other words, in the matrix model context, one has to consider glueball dynamics where one does not expect any in the standard gauge theory context.

Then, what about other low rank groups, such as $U(0)$, $SO(0)$ and $SO(2)$, where one does not consider a strongly coupled glueball dynamics in the standard gauge theory context? Should we include glueball fields also for them in the matrix model context, or not? If so (or if not), why? Clearly, these questions cannot be answered within the standard gauge theory framework; we should go back and investigate the string theory realization of these gauge theories, on which the Dijkgraaf–Vafa matrix model conjecture is based.

To answer these questions, in Chapter 4, we consider the string theory realization (geometric engineering) of supersymmetric $U(N)$, $SO(N)$, and $Sp(N)$ gauge theories with various two-index tensor matter fields and added tree-level superpotential, for general breaking patterns of the gauge group. These theories are realized in type IIB superstring theory compactified on certain noncompact Calabi–Yau 3-folds, as the world-volume theory on the D5-branes which wrap

compact 2-cycles in the Calabi–Yau space and fill four dimensional Minkowski spacetime. At low energies these gauge theories confine, developing a nonzero expectation value of the glueball field. This confinement transition is described in string theory by the geometric transition [8, 9, 10] where the 2-cycles in the Calabi–Yau space are blown down and 3-cycles are blown up instead. In string theory, the glueball field is interpreted as the size of these 3-cycles. This is an example of the geometry/gauge theory duality, where gravity in some background is dual to gauge theory in one less dimensions. Another related example of the duality is the celebrated AdS/CFT duality [32, 33, 34].

By closely studying the string theory physics near the blown up 3-cycles after the geometric transition, we clarify when glueball fields should be included and treated as dynamical fields, or rather set to zero. The resulting string theory prescription is the following: one should not include a glueball field for $U(0)$, $SO(0)$, $SO(2)$ while for all other gauge groups, including $U(1)$ and $Sp(0)$, one should consider an associated glueball field.

In particular, these string theory considerations give a clear physical explanation of the origin of the apparent discrepancy observed in [23] for $Sp(N)$ theory with antisymmetric tensor. Furthermore, we will give more examples in which this string theory prescription is crucial for obtaining the correct superpotential.

1.6 Adding flavors

So far, we discussed the string theory prescription when to include a glueball field, in theories with a matter field in the adjoint representation. It is a natural generalization to include matter fields in the fundamental and anti-fundamental representations, i.e., flavors.

For example, consider $\mathcal{N}=1$ $U(N)$ theory with an adjoint and N_f flavors. Classically, this theory has two kinds of vacua: the pseudo-confining vacua in which the gauge group breaks as $U(N) \to \prod_{i=1}^{K} U(N_i)$, $\sum_{i=1}^{K} N_i = N$, and the Higgs vacua in which $U(N) \to \prod_{i=1}^{K} U(N_i)$, $\sum_{i=1}^{K} N_i < N$. In the Higgs vacua, the total rank of the unbroken groups is reduced (Higgsed down), motivating the name of the vacua.

In the formalism of the generalized Konishi anomaly [35], this theory is described on a Riemann surface which is a double cover of the complex z-plane. On the z-plane, there are K cuts A_i, $i=1,\ldots,K$, corresponding to the K unbroken gauge group factors $U(N_i)$, $i=1,\ldots,K$. The glueball field S_i associated with the $U(N_i)$ group is related to a contour integral around the i-th cut, A_i. Moreover, there are N_f poles on the Riemann surface, one for each flavor. The position of a pole is determined by the mass of the corresponding flavor.

In this setup, the pseudo-confining vacua correspond to having all the N_f poles on the second sheet of the Riemann surface. On the other hand, the Higgs vacua is obtained as follows. One begins with the pseudo-confining vacua with all the poles on the second sheet, and starts varying the mass of a flavor smoothly. Then the position of the pole changes smoothly, and in particular, one can pass the pole through the cut A_i to the first sheet. In this process, the gauge group factor $U(N_i)$ becomes $U(N_i-1)$, i.e., it is Higgsed down. In this way, by passing poles on the second sheet through cuts to the first sheet, one can obtain the Higgs vacua. Different choices of the cuts through which the poles pass correspond to different Higgsing pattern.

Clearly, there should be a limit to this process of passing poles through a given cut, since as one passes poles one by one through the cut A_i associated with the unbroken $U(N_i)$ group, the group gets Higgsed down as $U(N_i) \to U(N_i-1) \to$

$U(N_i - 2) \to \cdots$, and one eventually ends up with a $U(0)$, which cannot be Higgsed down any further. Therefore, one expects that the cut shrinks as one passes more and more poles and it eventually closes up, when one has Higgsed down the $U(N_i)$ group completely. As mentioned above, the glueball S_i is related to the contour integral around the cut A_i. Therefore, when the $U(N_i)$ group has completely broken down and the cut A_i has closed up, the glueball vanishes: $S_i = 0$. This is consistent with the fact that there is no strongly coupled dynamics any more which gives rise to nonzero S_i.

One may think that the above picture is very reasonable and the $S_i \to 0$ limit should be describable in the matrix model framework in terms of the glueball S_i. However, that is not quite correct. Note that the Higgsing $U(N_i \neq 0) \to U(0)$ is not a smooth process, because in this process the number of massless $U(1)$ photons changes from one to zero, discontinuously. This is possible only by condensation of some charged massless particle [36, 37] which makes the photon massive by the Meissner effect and which is clearly missing in the glueball description.

In Chapter 5, we will closely study the above process of passing N_f poles through a cut associated with $U(N_c)$ group, and demonstrate that the cut indeed becomes arbitrarily small if one tries to pass too many poles through the cut. Furthermore, we will see that the situation where the cut has completely closed up, namely $S = 0$, is not describable in matrix model; we need some extra charged massless degree of freedom, as we discussed above. To find out what this massless field is, we will consider the string theory realization of the theory, and argue that this massless field should be a compact D3-brane which emanates fundamental strings. In addition, we will argue that the effect of this charged massless field can be taken care of by simply setting the glueball S_i to zero by hand in matrix

model, and demonstrate by explicit computation that the prescription indeed gives the correct superpotential obtained by the factorization method in gauge theory.

CHAPTER 2

$Sp(N)$ theory with antisymmetric tensor

With the aim of extending the gauge theory – matrix model connection to more general matter representations, we prove that for various two-index tensors of the classical gauge groups, the perturbative contributions to the glueball superpotential reduce to matrix integrals. Contributing diagrams consist of certain combinations of spheres, disks, and projective planes, which we evaluate to four and five loop order. In the case of $Sp(N)$ with antisymmetric matter, independent results are obtained by computing the nonperturbative superpotential for $N = 4, 6$ and 8. Comparison with the Dijkgraaf–Vafa approach reveals agreement up to $N/2$ loops in matrix model perturbation theory, with disagreement setting in at $h = N/2 + 1$ loops, h being the dual Coxeter number. At this order, the glueball superfield S begins to obey nontrivial relations due to its underlying structure as a product of fermionic superfields. We therefore find a relatively simple example of an $\mathcal{N} = 1$ gauge theory admitting a large N expansion, whose dynamically generated superpotential fails to be reproduced by the matrix model approach.

2.1 Introduction

The methods of Dijkgraaf and Vafa [5, 6, 7] represent a potentially powerful approach to obtaining nonperturbative results in a wide class of supersymmet-

ric gauge theories. Their original conjecture consists of two parts. First, that holomorphic physics is captured by an effective superpotential for a glueball superfield, with nonperturbative effects included via the Veneziano–Yankielowicz superpotential [38]. Second, that the Feynman diagrams contributing to the perturbative part of the glueball superpotential reduce to matrix model diagrams.

The second part of the conjecture has been proven for a few choices of matter fields and gauge groups, namely $U(N)$ with adjoint [13, 17] and fundamental [18] matter, and $SO/Sp(N)$ with adjoint matter [14, 39, 40]. Combining this with the first part of the conjecture has then been shown to reproduce known gauge theory results. Some examples of "exotic" tree-level superpotentials have also been considered successfully, such as multiple trace [41] and baryonic [42, 99, 100] interactions.

One naturally wonders how far this can be pushed. Generic $\mathcal{N} = 1$ theories possess intricate dynamically generated superpotentials which are difficult or (nearly) impossible to obtain by traditional means, and so a systematic method for computing them would be most welcome. The promise of the DV approach is that these perhaps can be obtained to any desired order by evaluating matrix integrals. With this in mind, we will demonstrate the reduction to matrix integrals for some new matter representations. Comparing with known gauge theory results will turn out to illustrate some apparent limitations of the DV approach.

In particular, it is straightforward to generalize the results of [13, 14] to more general two-index tensors of $U(N)$ and $SO/Sp(N)$, with or without tracelessness conditions imposed. The relevant $0 + 0$ dimensional Feynman diagrams which one needs to compute consist of various spheres, disks and projective planes, and disconnected sums of these. We evaluate these to five-loop order.

For comparison with gauge theory we focus on the particular case of $Sp(N)$[1] with an antisymmetric tensor chiral superfield. The dynamically generated superpotentials for such theories are highly nontrivial, and cannot be obtained via the "integrating in" approach of [43]. Furthermore, the results display no simple pattern in N. Nevertheless, a method is known for computing these superpotentials on a case-by-case basis [21, 22]. Results for $Sp(4)$ and $Sp(6)$ were obtained in [21, 22], and here we extend this to $Sp(8)$ as well (partial results for $Sp(8)$ appear in [22]). We believe that these examples illustrate the main features of generic $\mathcal{N} = 1$ superpotentials, and so are a good testing ground for the DV approach.

For our $Sp(N)$ examples, we will demonstrate agreement between gauge theory and the DV approach up to $N/2$ loops in perturbation theory, with a disagreement setting in at $N/2 + 1$ loops. In terms of the glueball superpotential, we thus find a disagreement at order S^h, where $h = N/2 + 1$ is the dual Coxeter number of $Sp(N)$.

The fact that discrepancies set in at order S^h is not a surprise, for it is at this order that S begins to obey relations due to its being a product of two fermionic superfields [17, 44]. Furthermore, at this order contributions to the effective action for W_α of the schematic form $\mathrm{Tr}(W_\alpha)^{2h}$ can be reexpressed in terms of lower traces, including S^h. Unfortunately, it is not clear how to ascertain these relations *a priori*, since they receive corrections from nonperturbative effects (see [44] for a recent discussion). These complications do not arise for theories with purely adjoint matter, since the results are known to have a simple pattern in N, and so N can be formally taken to infinity to avoid having to deal with any relations involving the S's. But in the more generic case, it seems that additional

[1]Our convention for $Sp(N)$ is such that N is an even integer, and $Sp(2) \cong SU(2)$.

input is required to make progress at h loops and beyond.

The remainder of this paper is organized as follows. In section 2.2 we isolate the field theory diagrams that contribute to the glueball superpotential, derive the reduction of these diagrams to those of a matrix model, and discuss their computation. These results are used in section 2.3 to derive effective superpotentials for $Sp(N)$ with matter in the antisymmetric tensor representation. In section 2.4 we state the corresponding results derived from a nonperturbative superpotential for these theories. Comparison reveals a discrepancy, which we discuss in section 2.5. Appendix 2.A gives more details on diagram calculations; appendix 2.B collects results from matrix model perturbation theory; and appendix 2.C concerns the computation of dynamically generated superpotentials for the $Sp(N)$ theories.

2.2 Reduction to matrix model

In this section we will extend the results of [13, 14] to include the following matter representations:

- $U(N)$ adjoint.
- $SU(N)$ adjoint.
- $SO(N)$ antisymmetric tensor
- $SO(N)$ symmetric tensor, traceless or traceful.
- $Sp(N)$ symmetric tensor.
- $Sp(N)$ antisymmetric tensor, traceless or traceful.

We will use Φ_{ij} to denote the matter superfield. In the case of $Sp(N)$, Φ_{ij} is defined as

$$\Phi = \begin{cases} SJ & S_{ij}\text{: symmetric tensor,} \\ AJ & A_{ij}\text{: antisymmetric tensor.} \end{cases} \tag{2.2.1}$$

Here J is the invariant antisymmetric tensor of $Sp(N)$, namely

$$J_{ij} = \begin{pmatrix} 0 & \mathbb{1}_{N/2} \\ -\mathbb{1}_{N/2} & 0 \end{pmatrix} . \tag{2.2.2}$$

The tracelessness of the Sp antisymmetric tensor is defined with respect to this J, i.e., by $\mathrm{Tr}[AJ] = 0$.

The fact that allows us to treat the above cases in parallel to those considered in [13, 14] is that gauge transformations act by commutation, $\delta_\Lambda \Phi \sim [\Lambda, \Phi]$. A separate analysis is needed for, say, $U(N)$ with a symmetric tensor.

2.2.1 Basic setup

Following [13], we consider a supersymmetric gauge theory with chiral superfield Φ and field strength $\mathcal{W}^\alpha$. Treating $\mathcal{W}^\alpha$ as a fixed background, we integrate out Φ to all orders in perturbation theory. We are interested in the part of the effective action which takes the form of a superpotential for the glueball superfield $S = \frac{1}{32\pi^2}\mathrm{Tr}[\mathcal{W}^\alpha \mathcal{W}_\alpha]$. In [13], using the superspace formalism, it was shown that this can be obtained from a simple action involving only chiral superfields:

$$S(\Phi) = \int d^4p\, d^2\pi \left[\frac{1}{2}\Phi(p^2 + \mathcal{W}^\alpha \pi_\alpha)\Phi + W_{\mathrm{tree}}(\Phi)\right] . \tag{2.2.3}$$

We choose the tree level superpotential to be

$$W_{\mathrm{tree}} = \frac{m}{2}\mathrm{Tr}[\Phi^2] + \text{interactions} , \tag{2.2.4}$$

where the interactions are single trace terms, and include the mass in the propagator:

$$\frac{1}{p^2+m+\mathcal{W}^\alpha\pi_\alpha}\ . \tag{2.2.5}$$

Actually, we have to be a little more precise here. Displaying all indices, we can write the quadratic action as

$$\frac{1}{2}\int d^4p\, d^2\pi\, \Phi_{ji} G^{-1}_{ijkl} \Phi_{kl} \tag{2.2.6}$$

with

$$G^{-1}_{ijkl} = \left[(p^2+m)\delta_{im}\delta_{jn} + (\mathcal{W}^\alpha)_{ijmn}\pi_\alpha\right] P_{mnkl}\ . \tag{2.2.7}$$

Here the P's are projection operators appropriate for the gauge group and matter representation under consideration:

$$P_{ijkl} = \begin{cases} \delta_{ik}\delta_{jl} & U(N) \text{ adjoint,} \\ \delta_{ik}\delta_{jl} - \frac{1}{N}\delta_{ij}\delta_{kl} & SU(N) \text{ adjoint,} \\ \frac{1}{2}(\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}) & SO(N) \text{ antisymmetric,} \\ \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) & SO(N) \text{ traceful symmetric,} \\ \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \frac{2}{N}\delta_{ij}\delta_{kl}) & SO(N) \text{ traceless symmetric,} \\ \frac{1}{2}(\delta_{ik}\delta_{jl} - J_{il}J_{jk}) & Sp(N) \text{ symmetric,} \\ \frac{1}{2}(\delta_{ik}\delta_{jl} + J_{il}J_{jk}) & Sp(N) \text{ traceful antisymmetric,} \\ \frac{1}{2}\left(\delta_{ik}\delta_{jl} + J_{il}J_{jk} - \frac{2}{N}\delta_{ij}\delta_{kl}\right) & Sp(N) \text{ traceless antisymmetric.} \end{cases} \tag{2.2.8}$$

The propagator is then given by the inverse of G^{-1} in the subspace spanned by P:

$$\langle \Phi_{ji}\Phi_{kl}\rangle = \left[\frac{P}{p^2+m+\mathcal{W}^\alpha\pi_\alpha}\right]_{ijkl} = \left[\int_0^\infty ds\ e^{-s(p^2+m+\mathcal{W}^\alpha\pi_\alpha)}P\right]_{ijkl}\ . \tag{2.2.9}$$

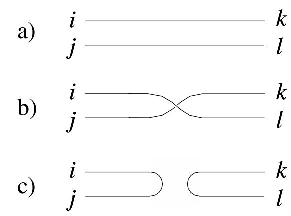


Figure 2.1: Propagators. a) untwisted; b) twisted; c) disconnected

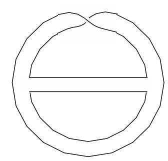


Figure 2.2: Typical diagram

Our rule for multiplying four-index objects is $(AB)_{ijkl} = \sum_{mn} A_{ijmn} B_{mnkl}$. The fact that gauge transformations act by commutation means that we can write

$$(\mathcal{W}^\alpha)_{ijkl} = (\mathcal{W}^\alpha)_{ik}\delta_{jl} - (\mathcal{W}^\alpha)_{lj}\delta_{ik} \ , \tag{2.2.10}$$

where on the right hand side $(\mathcal{W}^\alpha)_{ij}$ are field strengths in the defining representation of the gauge group.

2.2.2 Diagrammatics

The presence of the three sorts of terms in the projection operators (2.2.8) means that in double line notation we have three types of propagators, displayed in Fig. 2.1. Note in particular the disconnected propagator, which allows us to draw Feynman diagrams which have disconnected components in index space (All diagrams are connected in momentum space since we are computing the free energy). A typical diagram involving cubic interactions is shown in Fig. 2.2. Since we are computing the superpotential for S, we include either zero or two

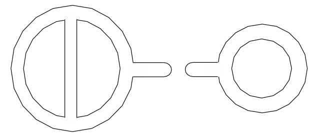


Figure 2.3: Contributing diagram

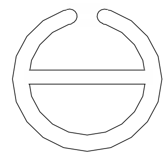


Figure 2.4: Non-contributing diagram

insertions of $\mathcal{W}^\alpha$ on each index loop.[2] We will now prove that the diagrams which contribute are those consisting of some number of sphere, disk, and projective plane components. Furthermore, the total number of disconnected components must be one greater than the number of disconnected propagators. Fig. 2.3 is an example of a contributing diagram, while Fig. 2.4 is a diagram which does not contribute.

The proof is similar to that given in [13, 14], so we mainly focus on the effect of the new disconnected propagator. In double line notation we associate each Feynman diagram to a two-dimensional surface. Let F by the number of faces (index loops); P be the number of edges; and V be the number of vertices. The Feynman diagram also has some number L of momentum loops. Euler's theorem tells us that

$$F = P - V + \chi \ , \tag{2.2.11}$$

[2] Note that we are explicitly *not* including the contributions coming from more than two $\mathcal{W}^\alpha$'s on an index loop, even if for a particular N these can be expressed in terms of S's. We will come back to this point in section 5.

where $\chi_{S^2} = 2$, $\chi_{D_2} = \chi_{\mathbb{RP}^2} = 1$. We also have the relation

$$F = L - 1 + \chi \ . \tag{2.2.12}$$

In a diagram with L loops we need to bring down L powers of S to saturate the fermion integrals, and we allow at most one S per index loop. Therefore, for a graph to be nonvanishing we need $F \geq L$. Graphs on S^2 with no disconnected propagators have $F = L + 1$, and those on $\mathbb{RP}^2$ have $F = L$.

To proceed we will make use of the following operation. Considering some diagram D that includes some number of disconnected propagators. To each D we associate a diagram $\tilde{D}$, obtained by replacing each disconnected propagator of D by an untwisted propagator. Each $\widetilde{D}$ diagram thus consists of a single connected component. D and $\widetilde{D}$ have the same values of L and V, but can have different values of F, P, and χ. We use $\widetilde{F}$, $\widetilde{P}$, and $\widetilde{\chi}$ to denote the number of faces, edges, and the Euler number of $\widetilde{D}$. To see which diagrams can contribute we consider various cases.

Case 1: D has no disconnected propagators, so $\widetilde{D} = D$. This case reduces to that of [13, 14], and so we know that only S^2 and $\mathbb{RP}^2$ graphs contribute (since no D_2 graphs can arise without disconnected propagators, these are the only graphs for which $F \geq L$.)

The remaining cases to consider are those for which we have at least one disconnected propagator.

Case 2: $\chi = \widetilde{\chi} \leq 1$

In this case $\widetilde{F} \leq L$, from (2.2.12). Each time we take a disconnected propagator and replace it by an untwisted propagator we are increasing P by 1 but keeping L unchanged. Therefore, from (2.2.11), this operation increases F by 1. So we see that in this case $F < L$. This means that the diagram D does not contribute.

Case 3: $\chi = \widetilde{\chi} = 2$

$\widetilde{D}$ has $\widetilde{F} = L + 1$. In this case, if D has a single disconnected propagator, we will have $F = L$, and so the diagram might seem to contribute. But we will now show that the fermion determinant vanishes for such diagrams.

We follow the conventions of [14], where the reader is referred for more details. The fermion contribution is proportional to $[\det N(s)]^2$ where

$$N(s)_{ma} = \sum_i s_i K^T_{mi} L_{ia} \ . \tag{2.2.13}$$

Here, i labels propagators; m labels "active" index loops on which we insert an S; and a labels momentum loops. In the present case, since $F = L$, all index loops are active and so N is a square matrix. To show that the determinant vanishes, we will show that the rectangular matrix

$$s_i K_{im} \tag{2.2.14}$$

has a nontrivial kernel.

Recall the definition of K_{im}. For each oriented propagator labeled by i, the mth index loop can do one of three things: 1) coincide and be parallel, giving $K_{im} = 1$; 2) coincide and be anti-parallel, giving $K_{im} = -1$; 3) not coincide, giving $K_{im} = 0$. Consider K_{im} acting on the vector b_m whose components are all equal to 1. It should be clear that

$$\sum_m K_{im} b_m = 1 - 1 = 0 \ . \tag{2.2.15}$$

The intuitive way to think about this is that b_m are the index loop momenta and $\sum_m K_{im} b_m$ are the propagator momenta. By setting all index loop momenta equal, one makes all propagator momenta vanish, and this corresponds to an element of the kernel of (2.2.14). This finally implies $\det N(s) = 0$, which is what we wanted to show.

Case 4: $\chi \neq \widetilde{\chi}$

This can only happen when D has two or more disconnected components. In this case, $\chi > \widetilde{\chi}$ and so (2.2.11) still allows $F \geq L$ even when $P < \widetilde{P}$.

In order to have a nonvanishing fermion integral, each component of D must have $F \geq L$, so each component must be an S^2, a D_2, or an $\mathbb{RP}^2$. Suppose D has N_{S^2} S^2 components, N_{D_2} D_2 components, and $N_{\mathbb{RP}^2}$ $\mathbb{RP}^2$ components, so that

$$\chi = 2N_{S^2} + N_{\mathbb{RP}^2} + N_{D_2} \ . \tag{2.2.16}$$

Next consider the relation between P and $\widetilde{P}$. The number of disconnected propagators must be at least the number of disconnected components of D minus one, so

$$P = \widetilde{P} - (N_{S^2} + N_{\mathbb{RP}^2} + N_{D_2} - 1) - a = \widetilde{P} + 1 + N_{S^2} - \chi - a \ , \tag{2.2.17}$$

where a is a nonnegative integer. Now use

$$F = P - V + \chi = \widetilde{P} + 1 + N_{S^2} - V - a \ . \tag{2.2.18}$$

$\widetilde{D}$ satisfies

$$\widetilde{P} + 1 - V = L \ , \tag{2.2.19}$$

so we get

$$F = L + N_{S^2} - a \ . \tag{2.2.20}$$

Now, in order to have a nonzero fermion determinant we need to have at least one inactive index loop (no $\mathcal{W}^a$ insertions) per S^2 component after choosing L active index loops. In other words, a nonvanishing fermion determinant requires

$$F \geq L + N_{S^2} \ . \tag{2.2.21}$$

Putting these two conditions together, we clearly need $a = 0$. This says that the number of disconnected propagators in D must be precisely equal to the number of disconnected components of D minus one.

Summary: Diagrams which contribute to the glueball superpotential have any number of disconnected S^2, D_2, and $\mathbb{RP}^2$ components. The number of disconnected propagators must be one less than the number of disconnected components.[3]

2.2.3 Computation of diagrams

Now that we have isolated the class of diagrams which contribute to the glueball superpotential, we turn to their computation. This turns out to be a simple extension of what is already known. In particular, the contribution from a general disconnected diagram is simply equal to an overall combinatorial factor times the product of the contributions of the individual components. This follows from the fact that, for the diagrams we are considering, the disconnected propagators carry vanishing momentum, so the diagrams are actually disconnected in both momentum space and index space.

Next, we observe that the stubs from the disconnected propagators can be neglected in the computation; it is easily checked that the sum over $\mathcal{W}^\alpha$ insertions on the stubs gives zero due to the minus sign in (2.2.10).

So we just need rules for treating each component individually, and then we multiply the contributions together to get the total diagram. The rules consist of relating the gauge theory contribution to a corresponding matrix contribution.

[3] It is easy to convince oneself that disconnected diagrams will therefore never contribute in theories with only even powers in the tree level superpotential, thus giving the same glueball superpotential in the traceful and traceless cases.

The cases of interest are:

S^2 **components:** From the work of [13], we know that if $N^{L+1}F^{(L)}_{S^2}(g_k)$ is the contribution in the matrix model diagram from an L loop S^2 graph, then the contribution in the gauge theory diagram is

$$W^{(L)}_{S^2}(S, g_k) = (L+1)NS^L F^{(L)}_{S^2}(g_k) \ . \tag{2.2.22}$$

The prefactor $(L+1)N$ comes from the choice of, and trace over, a single inactive index loop.

$\mathbb{RP}^2$ **components:** From the work of [14], we know that if $NF^{(L)}_{\mathbb{RP}^2}(g_k)$ is the contribution in the matrix model diagram from an L loop $\mathbb{RP}^2$ graph, then the contribution in the gauge theory diagram is

$$W^{(L)}_{\mathbb{RP}^2}(S, g_k) = \pm 4S^L F^{(L)}_{\mathbb{RP}^2}(g_k) \ . \tag{2.2.23}$$

The prefactor of ± 4 comes from the fermion determinant, and is equal to $+4(-4)$ for symmetric(antisymmetric) tensors.

D_2 **components:** These have $L = 0$ and hence no $\mathcal{W}^a$ insertions. So if the contribution to the matrix model is $NF^{(L)}_{D_2}(g_k)$ then

$$W^{(L)}_{D_2}(g_k) = NF^{(L)}_{D_2}(g_k) \ . \tag{2.2.24}$$

With the above rules in hand, it is a simple matter to convert a given matrix model Feynman diagram into a contribution to the glueball superpotential. The example given in Appendix 2.A should help to clarify this. We should emphasize that the above procedure must by done diagram by diagram — there is no obvious way to directly relate the entire glueball superpotential to the matrix model free energy; the situation is similar to [41] in this respect.

2.3 Results from matrix integrals

The considerations thus far apply to any single trace, polynomial, tree level superpotential. We now restrict attention to cubic interactions,

$$W_{\text{tree}} = \frac{m}{2}\mathrm{Tr}\,\Phi^2 + \frac{g}{3}\mathrm{Tr}\,\Phi^3 \ , \tag{2.3.1}$$

(which are of course trivial in the case of SO/Sp with adjoint matter.) In Appendix 2.B we collect our matrix model results for the various matter representations. In this section we focus on two particular cases, which will be compared to gauge theory results in the next section.

2.3.1 $Sp(N)$ with traceful antisymmetric matter

The perturbative part of the glueball superpotential for $Sp(N)$ with traceful antisymmetric matter is

$$\begin{aligned} W^{\text{pert}}_{\text{traceful}}(S,\alpha) = &\,(-N+3)\,\alpha S^2 + \left(-\frac{16}{3}N + \frac{59}{3}\right)\alpha^2 S^3 \\ &+ \left(-\frac{140}{3}N + 197\right)\alpha^3 S^4 + \left(-512N + \frac{4775}{2}\right)\alpha^4 S^5 + \cdots \end{aligned} \tag{2.3.2}$$

with

$$\alpha \equiv \frac{g^2}{2m^3} \ . \tag{2.3.3}$$

In terms of diagrams, (2.3.2) represents the contribution from $2,3,4$ and 5 loops. According to the DV conjecture, the full glueball superpotential is then $W^{\text{eff}} = W^{VY} + W^{\text{pert}}$, where W^{VY} is the Veneziano–Yankielowicz superpotential:

$$W^{VY} = (N/2+1)S[1-\log(S/\Lambda^3)] \ . \tag{2.3.4}$$

We are now instructed to extremize W^{eff} with respect to S and substitute back in. We call the result W^{DV}. Working in a power series in g, we obtain

$$\begin{aligned} W^{\text{DV}}_{\text{traceful}}(\Lambda, m, g) = (N/2+1)\Lambda^3 \Bigg[1 - \frac{2(N-3)}{N+2}\Lambda^3\alpha - \frac{2(4N^2+45N-226)}{3(N+2)^2}\Lambda^6\alpha^2 \\ - \frac{2(12N^3+293N^2+368N-8340)}{3(N+2)^3}\Lambda^9\alpha^3 \\ - \frac{96N^4+3803N^3+25868N^2-85092N-744768}{3(N+2)^4}\Lambda^{12}\alpha^4 - \cdots \Bigg] . \quad (2.3.5) \end{aligned}$$

For $N = 4, 6, 8$, this yields

$$\begin{aligned} W^{\text{DV},Sp(4)}_{\text{traceful}}(\Lambda, \alpha) &= 3\Lambda^3 - \Lambda^6\alpha - \Lambda^9\alpha^2 - \frac{353}{27}\Lambda^{12}\alpha^3 - \frac{25205}{81}\Lambda^{15}\alpha^4 - \cdots , \\ W^{\text{DV},Sp(6)}_{\text{traceful}}(\Lambda, \alpha) &= 4\Lambda^3 - 3\Lambda^6\alpha - \frac{47}{6}\Lambda^9\alpha^2 - \frac{73}{2}\Lambda^{12}\alpha^3 - \frac{6477}{32}\Lambda^{15}\alpha^4 - \cdots , \\ W^{\text{DV},Sp(8)}_{\text{traceful}}(\Lambda, \alpha) &= 5\Lambda^3 - 5\Lambda^6\alpha - 13\Lambda^9\alpha^2 - 65\Lambda^{12}\alpha^3 - \frac{2142}{5}\Lambda^{15}\alpha^4 - \cdots . \end{aligned} \quad (2.3.6)$$

2.3.2 $Sp(N)$ with traceless antisymmetric matter

Including the contribution from the disconnected propagator, the perturbative part of the glueball superpotential for $Sp(N)$ with traceless antisymmetric matter is

$$\begin{aligned} W^{\text{pert}}_{\text{traceless}}(S, \alpha) = \left(-1 + \frac{4}{N}\right)\alpha S^2 + \left(-\frac{1}{3} - \frac{8}{N} + \frac{160}{3N^2}\right)\alpha^2 S^3 \\ + \left(-\frac{1}{3} - \frac{12}{N} - \frac{256}{3N^2} + \frac{3584}{3N^3}\right)\alpha^3 S^4 + \cdots . \quad (2.3.7) \end{aligned}$$

The presence of many disconnected diagrams makes this case more complicated than the traceful case, and we have correspondingly worked to one lower order than in (2.3.2).

Adding the Veneziano–Yankielowicz superpotential and integrating out the

glueball superfield, we obtain

$$W^{\mathrm{DV}}_{\mathrm{traceless}}(\Lambda, m, g) = (N/2+1)\Lambda^3 \Bigg[1 - \frac{2(N-4)}{N(N+2)}\Lambda^3\alpha - \frac{2(N^3+14N^2-16N-512)}{3N^2(N+2)^2}\Lambda^6\alpha^2 - \frac{2(N+8)^2(N^3+12N^2-52N-528)}{3N^3(N+2)^3}\Lambda^9\alpha^3 - \cdots \Bigg] . \quad (2.3.8)$$

This yields

$$\begin{aligned} W^{\mathrm{DV},Sp(4)}_{\mathrm{traceless}}(\Lambda,\alpha) &= 3\Lambda^3 + \Lambda^9\alpha^2 + 10\Lambda^{12}\alpha^3 + \cdots \ , \\ W^{\mathrm{DV},Sp(6)}_{\mathrm{traceless}}(\Lambda,\alpha) &= 4\Lambda^3 - \frac{1}{3}\Lambda^6\alpha - \frac{7}{54}\Lambda^9\alpha^2 + \frac{49}{54}\Lambda^{12}\alpha^3 + \cdots \ , \\ W^{\mathrm{DV},Sp(8)}_{\mathrm{traceless}}(\Lambda,\alpha) &= 5\Lambda^3 - \frac{1}{2}\Lambda^6\alpha - \frac{2}{5}\Lambda^9\alpha^2 - \frac{14}{25}\Lambda^{12}\alpha^3 - \cdots \ . \end{aligned} \quad (2.3.9)$$

2.4 Gauge theory example: $Sp(N)$ with antisymmetric matter

Dynamically generated superpotentials can be determined for $\mathcal{N}=1$ theories with gauge group $Sp(N)$ and a chiral superfield A_{ij} in the antisymmetric tensor representation. The general procedure was given in [21, 22], and is reviewed in Appendix C. Since these superpotentials cannot be obtained by the integrating in procedure of [43], they are more difficult to establish, and the results are correspondingly more involved, than more familiar examples. A separate computation is required for each N, and the results display no obvious pattern in N. $N=4$ is a simple special case (since $Sp(4) \cong SO(5)$ and $A_{ij} \cong$ vector); $N=6$ was worked out in [21, 22], and in Appendix C we extend this to $Sp(8)$ ([22] gives the result for $Sp(8)$ with some additional fundamentals, which need to be integrated out for our purposes). In this section we state the results, and integrate out A_{ij} to obtain formulas that we can compare with the DV approach.

The moduli space of the classical theory is parameterized by the gauge invariant operators

$$O_n = \mathrm{Tr}[(AJ)^n], \qquad n = 1, 2, \ldots, N/2 \ , \tag{2.4.1}$$

the upper bound coming from the characteristic equation of the matrix AJ.

From the gauge theory point of view, it is natural to demand tracelessness, and this will be denoted by a tilde: $\mathrm{Tr}[\widetilde{A}J] = 0$,

$$\widetilde{O}_n = \mathrm{Tr}[(\widetilde{A}J)^n], \qquad n = 2, \ldots, N/2 \ . \tag{2.4.2}$$

In comparing with the DV approach, we will consider both the traceless and traceful cases.

2.4.1 Traceless case

The $Sp(4)$ and $Sp(6)$ dynamical superpotentials for these fields are [21, 22]:

$$W_{\rm dyn}^{Sp(4)} = \frac{2{\Lambda_0}^4}{\widetilde{O}_2^{1/2}} \ , \tag{2.4.3}$$

$$W_{\rm dyn}^{Sp(6)} = \frac{4{\Lambda_0}^5}{\widetilde{O}_2[(\sqrt{R}+\sqrt{R+1})^{2/3}+(\sqrt{R}+\sqrt{R+1})^{-2/3}-1]} \ , \tag{2.4.4}$$

with $R = -12{\widetilde{O}_3}^2/{\widetilde{O}_2}^3$.

Also, as derived in Appendix 2.C, the $Sp(8)$ superpotential is

$$\begin{aligned} W_{\rm dyn}^{Sp(8)} = \frac{6\sqrt{2}\,\Lambda_0^6}{\widetilde{O}_2^{3/2}}\Big[&-36R_4 + 144b^2R_4 + 288cR_4 + 8R_3^2 \\ &+ 192bcR_3 + 1152b^2c^2 - 36b^2 - 72c + 9\Big]^{-1} \ , \end{aligned} \tag{2.4.5}$$

where $R_3 \equiv \widetilde{O}_3/{\widetilde{O}_2}^{3/2}$, $R_4 \equiv \widetilde{O}_4/{\widetilde{O}_2}^2$, and b and c are determined by

$$\begin{aligned} 12R_4 + 16bR_3 - 192b^2c + 24b^2 + 96c^2 - 3 = 0 \ , \\ 12bR_4 + 8b^2R_3 + 8R_3c - 96bc^2 + 24bc - 3b = 0 \ . \end{aligned} \tag{2.4.6}$$

We choose the root which gives $R_3 = 0$ as the solution of the F-flatness condition.

Now let us integrate out the antisymmetric matter. We add the tree level superpotential

$$W_{\text{tree}} = \frac{m}{2}\widetilde{O}_2 + \frac{g}{3}\widetilde{O}_3 \tag{2.4.7}$$

to the dynamical part, solve the F-flatness equations, and substitute back in. We do this perturbatively in g, and obtain

$$\begin{aligned}
W^{\text{gt},Sp(4)}_{\text{traceless}} &= 3\Lambda^3, \\
W^{\text{gt},Sp(6)}_{\text{traceless}} &= 4\Lambda^3 - \frac{1}{3}\Lambda^6\alpha - \frac{7}{54}\Lambda^9\alpha^2 - \frac{5}{54}\Lambda^{12}\alpha^3 - \frac{221}{2592}\Lambda^{15}\alpha^4 - \cdots\ , \\
W^{\text{gt},Sp(8)}_{\text{traceless}} &= 5\Lambda^3 - \frac{1}{2}\Lambda^6\alpha - \frac{2}{5}\Lambda^9\alpha^2 - \frac{14}{25}\Lambda^{12}\alpha^3 - \Lambda^{15}\alpha^4 - \cdots\ ,
\end{aligned} \tag{2.4.8}$$

where α is defined in (2.3.3), and the low-energy scales are defined from the usual matching conditions as

$$\begin{aligned}
Sp(4):&\quad \Lambda^9 = (\frac{m}{2})\Lambda_0^8\ , \\
Sp(6):&\quad \Lambda^6 = (\frac{m}{2})\Lambda_0^5\ , \\
Sp(8):&\quad \Lambda^{15} = (\frac{m}{2})^3\Lambda_0^{12}\ .
\end{aligned} \tag{2.4.9}$$

2.4.2 Traceful case

For $Sp(N)$ theory with a traceful antisymmetric tensor A_{ij}, we separate out the trace part as

$$A_{ij} = \widetilde{A}_{ij} - \frac{1}{N}J_{ij}\phi, \qquad \text{Tr}[\widetilde{A}J] = 0, \qquad \text{Tr}[AJ] = \phi\ . \tag{2.4.10}$$

$\widetilde{O}_n$ are related to their traceful counterparts $O_n \equiv \text{Tr}[(AJ)^n]$ by

$$O_2 = \widetilde{O}_2 + \frac{\phi^2}{N}, \qquad O_3 = \widetilde{O}_3 + \frac{3}{N}\widetilde{O}_2\phi + \frac{1}{N^2}\phi^3. \tag{2.4.11}$$

The dynamical superpotential of this traceful theory is the same as the traceless theory, since ϕ has its own $U(1)_\phi$ charge and hence cannot enter in W_{dyn}.

Integrating out $\widetilde{A}_{ij}$ and ϕ in the presence of the tree level superpotential

$$W_{\text{tree}} = \frac{m}{2}O_2 + \frac{g}{3}O_3, \tag{2.4.12}$$

we obtain

$$\begin{aligned}
W^{\text{gt},Sp(4)}_{\text{traceful}} &= 3\Lambda^3 - \Lambda^6\alpha - 2\Lambda^9\alpha^2 - \frac{187}{27}\Lambda^{12}\alpha^3 - \frac{2470}{27}\Lambda^{15}\alpha^4 - \cdots \ , \\
W^{\text{gt},Sp(6)}_{\text{traceful}} &= 4\Lambda^3 - 3\Lambda^6\alpha - \frac{47}{6}\Lambda^9\alpha^2 - \frac{75}{2}\Lambda^{12}\alpha^3 - \frac{7437}{32}\Lambda^{15}\alpha^4 - \cdots \ , \\
W^{\text{gt},Sp(8)}_{\text{traceful}} &= 5\Lambda^3 - 5\Lambda^6\alpha - 13\Lambda^9\alpha^2 - 65\Lambda^{12}\alpha^3 - \frac{2147}{5}\Lambda^{15}\alpha^4 - \cdots \ .
\end{aligned} \tag{2.4.13}$$

2.5 Comparison and discussion

According to the general conjecture, we are supposed to compare (2.3.9) with (2.4.8), and (2.3.6) with (2.4.13). We write $\triangle W \equiv W^{\text{DV}} - W^{\text{gt}}$, and find

$$\begin{aligned}
\triangle W^{Sp(4)}_{\text{traceless}} &= 0\cdot\Lambda^6\alpha + \Lambda^9\alpha^2 + \cdots \ , \\
\triangle W^{Sp(6)}_{\text{traceless}} &= 0\cdot\Lambda^6\alpha + 0\cdot\Lambda^9\alpha^2 + \Lambda^{12}\alpha^3 + \cdots \ , \\
\triangle W^{Sp(8)}_{\text{traceless}} &= 0\cdot\Lambda^6\alpha + 0\cdot\Lambda^9\alpha^2 + 0\cdot\Lambda^{12}\alpha^3 + \mathcal{O}(\Lambda^{15}\alpha^4) \ .
\end{aligned} \tag{2.5.1}$$

and

$$\begin{aligned}
\triangle W^{Sp(4)}_{\text{traceful}} &= 0\cdot\Lambda^6\alpha + \Lambda^9\alpha^2 + \cdots \ , \\
\triangle W^{Sp(6)}_{\text{traceful}} &= 0\cdot\Lambda^6\alpha + 0\cdot\Lambda^9\alpha^2 + \Lambda^{12}\alpha^3 + \cdots \ , \\
\triangle W^{Sp(8)}_{\text{traceful}} &= 0\cdot\Lambda^6\alpha + 0\cdot\Lambda^9\alpha^2 + 0\cdot\Lambda^{12}\alpha^3 + \Lambda^{15}\alpha^4 + \cdots \ .
\end{aligned} \tag{2.5.2}$$

We have indicated the terms that canceled nontrivially by including them with a coefficient of zero. From these examples, we see that a disagreement sets in at order $(\Lambda^3)^h\alpha^{h-1}$, where $h = N/2+1$ is the dual Coxeter number. We also observe that the coefficient of the disagreement at this order is unity. We now discuss the implications of this result.

First, it is very unlikely that the discrepancy is due to a computational error, such as forgetting to include a diagram. This is apparent from the fact that

the mismatch arises at a different order in perturbation theory for different rank gauge groups. So adding a new contribution to the $Sp(4)$ result at order $\Lambda^9\alpha^2$, say, would generically destroy the agreement for $Sp(6)$ and $Sp(8)$ at this order. Instead, it is much more likely that our results indicate a breakdown of the underlying approach.

Let us return to the two basic elements of the DV conjecture. The first part asserts that the perturbative part of the glueball superpotential can be computed from matrix integrals, and the second part assumes that nonperturbative effects are captured by adding the Veneziano–Yankielowicz superpotential. We have proven the perturbative part of the conjecture for the relevant matter fields, but there is one subtlety which we have so far avoided but now must discuss.

In our perturbative computations we inserted no more than two $\mathcal{W}_\alpha$'s on any index loop, since we were interested in a superpotential for $S \sim \mathrm{Tr}\mathcal{W}^2$, and not in operators such as Tr $\mathcal{W}^{2n}$, $n > 1$. However, for a given gauge group, it may be possible to use Lie algebra identities to express such "unwanted" operators in terms of other operators, including S. Should we then include these new S terms along with our previous results?

This issue seems especially pertinent given that our discrepancy sets in at order S^h, which is when we begin to find nontrivial relations involving S due to its underlying structure as a product of the fermionic field $\mathcal{W}_\alpha$. For example, for $Sp(4)$ there are relations such as

$$\mathrm{Tr}[(\mathcal{W}^2)^3] = \frac{3}{4}\mathrm{Tr}[\mathcal{W}^2]\mathrm{Tr}[(\mathcal{W}^2)^2] - \frac{1}{8}(\mathrm{Tr}[\mathcal{W}^2])^3. \tag{2.5.3}$$

So a naive guess is that the discrepancy can be accounted for if we keep all contributions coming from more than two $\mathcal{W}_\alpha$'s on an index loop, and re-express the traces of the form $\mathrm{Tr}(\mathcal{W}_\alpha)^{2n}$ $(n \geq h)$ in favor of S using relations like (2.5.3), setting all traces to zero that are not re-expressible in terms of S. Such consider-

ations are indeed necessary in order to avoid getting nonsensical results in certain cases, e.g. antisymmetric matter for $Sp(2)$. Such a matter field is uncharged, and so should certainly contribute a vanishing result for the glueball superpotential, but this is seen only if we compute all the trace structures. We should emphasize that if we keep all of these contributions then perturbation theory will not reduce to matrix integrals since the Schwinger parameter dependence will not cancel; nevertheless we can try this procedure and see what we get.

In order to check if the above guess is correct, we took a Φ^{2p} interaction and evaluated the perturbative superpotential explicitly keeping all the traces. For this interaction, a discrepancy arises at the first order if we take $p > N$. Specifically, we considered a Φ^6 interaction in $Sp(4)$ with antisymmetric matter. After a tedious calculation, we have found that this does *not* account for the discrepancy.

It therefore seems more likely that the problem lies in the nonperturbative part of the conjecture, involving truncating to just S (and dropping other operators like $\mathrm{Tr}[(\mathcal{W}^2)^2]$), and just adding the Venziano–Yankielowicz superpotential. There is no solid motivation for this procedure beyond the fact that it seems to give sensible results in various cases. Our results indicate that for generic theories this recipe is valid only up to $(h-1)$ loops. On the other hand, the fact that our discrepancies arise in a very simple fashion — always with a coefficient of unity — suggests that perhaps there exists a way of generalizing the DV recipe to enable us to go to h loops and beyond.

Clearly, it is important to resolve these issues in order to determine the range of validity of the DV approach. One might have hoped that the approach would be useful for any $\mathcal{N}=1$ theory admitting a large N expansion. Our $Sp(N)$ theories are certainly in this class, and so seem to provide an explicit counterexample.

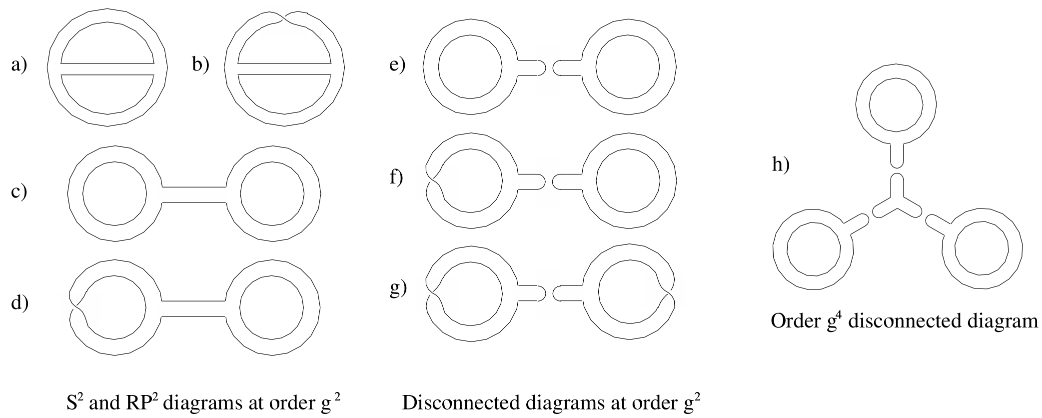


Figure 2.5: Diagrams for traceless tensor matter field

Appendix

2.A diagrammatics for traceless matter field

In this appendix, we sketch the diagrammatics for evaluating the perturbative glueball superpotential, focusing on the case with a traceless tensor. To be specific, we consider the cubic interaction below:

$$e^{-W^{\mathrm{pert}}(S)} = \int \mathcal{D}\Phi \, e^{-\int d^4x\, d^2\theta \, \mathrm{Tr}\left[-\frac{1}{2}\Phi(\partial^2 - i\mathcal{W}^\alpha D_\alpha)\Phi + \frac{m}{2}\Phi^2 + \frac{g}{3}\Phi^3\right]}. \tag{2.A.1}$$

Namely, we consider SO with traceless symmetric matter, or Sp with traceless antisymmetric matter.

At order g^2, there are four S^2 and $\mathbb{RP}^2$ diagrams without disconnected propagators that contribute, as shown in Fig. 2.5 a)–d). These can be evaluated by combinatorics. For a) and b) there are 6 ways to contract legs. For c) and d), on the other hand, there are 3^2 ways to contract legs and 2 choices for the middle propagator (untwisted or twisted). Since there are two loop momenta, the glueball $S \sim \mathcal{W}^\alpha \mathcal{W}_\alpha$ should be inserted in two index loops, and we may insert only up to one glueball on each index loop. For S^2 graphs a) and c), there are 3

ways to do so, and a trace over the remaining index loop contributes N. For $\mathbb{RP}^2$ graphs b) and d), there is only one way to insert the glueball, but the fermionic determinant gives an extra factor (± 4). The sign depends on the matter field under consideration see section 2.) Finally, for b) and d) there are respectively 3 and 2 ways to choose which propagator to twist. Therefore, the contributions are

$$\begin{aligned} f_{\mathrm{a}} &= 6 \cdot 3 \cdot NS^2, \quad f_{\mathrm{b}} = 6 \cdot 3 \cdot (\pm 4) S^2, \\ f_{\mathrm{c}} &= 3^2 \cdot 2 \cdot 3 \cdot NS^2, \quad f_{\mathrm{d}} = 3^2 \cdot 2 \cdot 2 \cdot (\pm 4) S^2. \end{aligned} \tag{2.A.2}$$

Including factors coming from propagators and coefficients from Taylor expansion, we obtain

$$W^{(2)}_{\mathrm{conn}} = -\frac{g^2}{(2m)^3} \frac{1}{2!\, 3^2} (f_{\mathrm{a}} + f_{\mathrm{b}} + f_{\mathrm{c}} + f_{\mathrm{d}}) = (-N \mp 3) \alpha S^2. \tag{2.A.3}$$

where we defined $\alpha = \frac{g^2}{2m^3}$ as before. This reproduces the first term of the traceful result (2.3.2).

For a traceless tensor, there are three additional diagrams e), f) and g), with disconnected propagators that give nonvanishing contributions.

These can be evaluated similarly to the connected ones. First, there are factors common to all three graphs; $(-2/N)$ from the disconnected propagator, and $3^2 = 9$ from the ways to contract legs. In addition, the particular graphs have the additional factors; e): $(2N)^2$ from the ways of inserting a glueball in one of two index loops in each S^2 component, and the trace on the remaining index loop. f): (± 4) from the fermionic determinant of the $\mathbb{RP}^2$ component, and $2N$ from the glueball insertion into the S^2 component. Also, there is the same contribution from the $S^2 \times \mathbb{RP}^2$ graph. g): $(\pm 4)^2$ from two $\mathbb{RP}^2$ components.

Altogether we obtain

$$W^{(2)}_{\text{disconn}} = -\frac{g^2}{(2m)^3}\frac{1}{2!\,3^2}\left(-\frac{2}{N}\right)\cdot 9\cdot[(2N)^2+2\cdot 2N\cdot(\pm 4)+(\pm 4)^2]S^2$$
$$= \frac{(2N\pm 4)^2}{4N}\alpha S^2. \tag{2.A.4}$$

Summing the connected and disconnected contributions, we obtain

$$W^{(2)}_{\text{conn}} + W^{(2)}_{\text{disconn}} = \left(\pm 1 + \frac{4}{N}\right)\alpha S^2 \tag{2.A.5}$$

which is the first term of the traceless result (2.3.7).

Higher order diagrams can be worked out in much the same way, although the number of diagrams increases rapidly. For the disconnected diagrams, we only have to consider the diagrams which are one-particle-reducible with respect to the disconnected propagator. Therefore, we basically just splice lower order diagrams with the disconnected propagator. The contribution is just the product of the contributions from the lower order pieces, multiplied by the ways to insert the disconnected propagator into them, and by $(-2/N)^n$ from the disconnected propagator itself. However, note that one should also consider diagrams such as h) of Fig. 2.5. In this case, the central D_2 piece contributes N from its index loop.

2.B Summary of results from perturbation theory

In this appendix we state our results for $W^{\text{pert}}(S,\alpha)$, the perturbative contributions to the glueball superpotential. These correspond to evaluating certain diagrams in the matrix model. We consider cubic interactions only,

$$W_{\text{tree}} = \frac{m}{2}\operatorname{Tr}\Phi^2 + \frac{g}{3}\operatorname{Tr}\Phi^3 \ , \tag{2.B.1}$$

which means that we will not consider the case of SO/Sp with adjoint matter. In any event, it is not necessary to compute the perturbative superpotential for the latter cases, since closed form expressions for even power interactions are already known [14, 39, 40]. The case of $U(N)$ with adjoint matter is also well known [45], but for convenience we include it in the list below.

For traceful matter fields, instead of evaluating individual Feynman diagrams, there exists a much simpler method for computing which we have used to obtain the results below. We can simply compute the matrix model free energy by computer for certain low values of N. Since S^2 and $\mathbb{RP}^2$ diagrams scale as N^{L+1} and N^L at L loops in perturbation theory, we can easily read off the S^2 and $\mathbb{RP}^2$ contributions to any desired order. For traceless fields things are not so simple, since the N dependence becomes more complicated, and certain diagrams must be discarded (as discussed in Section 2.)

We define

$$\alpha = \begin{cases} g^2/m^3 & U(N) \ , \\ 2g^2/m^3 & SO/Sp(N) \ . \end{cases} \tag{2.B.2}$$

2.B.1 $U(N)$ with adjoint matter

$$W^{\text{pert}}(S,\alpha) = N\frac{\partial F_{\chi=2}}{\partial S} \ , \qquad F_{\chi=2} = -\frac{S^2}{2}\sum_{k=1}^{\infty}\frac{(8\alpha S)^k}{(k+2)!}\frac{\Gamma\left(\frac{3k}{2}\right)}{\Gamma\left(\frac{k}{2}+1\right)} \ , \tag{2.B.3}$$

$$W^{\text{pert}}(S,\alpha) = -2N\alpha S^2 - \frac{32}{3}N\alpha^2 S^3 - \frac{280}{3}N\alpha^3 S^4 - 1024N\alpha^4 S^5 - \cdots \ . \tag{2.B.4}$$

2.B.2 $SU(N)$ with adjoint matter

$$W^{\text{pert}}(S,\alpha) = 0\cdot\alpha S^2 + 0\cdot\alpha^2 S^3 + 0\cdot\alpha^3 S^4 + \cdots \ . \tag{2.B.5}$$

2.B.3 $SO(N)$ with traceful symmetric matter

$$W^{\text{pert}}(S,\alpha) = -(N+3)\,\alpha S^2 - \left(\frac{16}{3}N + \frac{59}{3}\right)\alpha^2 S^3 - \left(\frac{140}{3}N + 197\right)\alpha^3 S^4 - \left(512N + \frac{4775}{2}\right)\alpha^4 S^5 - \cdots \tag{2.B.6}$$

2.B.4 $SO(N)$ with traceless symmetric matter

$$W^{\text{pert}}(S,\alpha) = \left(1 + \frac{4}{N}\right)\alpha S^2 + \left(\frac{1}{3} - \frac{8}{N} - \frac{160}{3N^2}\right)\alpha^2 S^3 + \left(\frac{1}{3} - \frac{12}{N} + \frac{256}{3N^2} + \frac{3584}{3N^3}\right)\alpha^3 S^4 + \cdots\ . \tag{2.B.7}$$

2.B.5 $Sp(N)$ with traceful antisymmetric matter

$$W^{\text{pert}}(S,\alpha) = (-N+3)\,\alpha S^2 + \left(-\frac{16}{3}N + \frac{59}{3}\right)\alpha^2 S^3 + \left(-\frac{140}{3}N + 197\right)\alpha^3 S^4 + \left(-512N + \frac{4775}{2}\right)\alpha^4 S^5 + \cdots\ . \tag{2.B.8}$$

2.B.6 $Sp(N)$ with traceless antisymmetric matter

$$W^{\text{pert}}(S,\alpha) = \left(-1 + \frac{4}{N}\right)\alpha S^2 + \left(-\frac{1}{3} - \frac{8}{N} + \frac{160}{3N^2}\right)\alpha^2 S^3 + \left(-\frac{1}{3} - \frac{12}{N} - \frac{256}{3N^2} + \frac{3584}{3N^3}\right)\alpha^3 S^4 + \cdots\ . \tag{2.B.9}$$

These results exhibit some remarkable cancellations. We find a vanishing result for $SU(N)$ with adjoint matter, and a cancellation of the terms linear in N for $Sp(N)$ with traceless antisymmetric matter. In both cases the cancellation seems to involve all the diagrams at a given order. We do not have a proof of cancellation beyond the order indicated; it would be nice to provide one and to better understand the significance of this fact.

2.C Gauge theory results

In [21, 22], a systematic method for determining the dynamical superpotential of the $Sp(N)$ gauge theory with a traceless antisymmetric matter $\widetilde{A}_{ab}$ was proposed. In this appendix, we briefly review the strategy, focusing on the $Sp(8)$ case.

First, we add to the theory $2N_F$ fundamentals Q_i. The moduli space of this enlarged $(N_{\widetilde{A}}, N_F)$ theory is parameterized by

$$\widetilde{O}_n = \mathrm{Tr}[(\widetilde{A}J)^n], \qquad n = 2, 3, \cdots, N/2 \tag{2.C.1}$$

as well as the antisymmetric matrices

$$\begin{aligned} M_{ij} &= Q_i^T J Q_j \ , \\ N_{ij} &= Q_i^T J \widetilde{A} J Q_j \ , \\ P_{ij} &= Q_i^T J (\widetilde{A}J)^2 Q_j \ , \\ &\cdots \\ R_{ij} &= Q_i^T J (\widetilde{A}J)^{k-1} Q_j \ . \end{aligned} \tag{2.C.2}$$

The basic observation is that for $N_F = 3$, symmetry and holomorphy considerations restrict the dynamical superpotential to be of the form

$$W^{\mathrm{dyn}}_{(1,3)} = \frac{\text{Some polynomial in } \widetilde{O}_n, M_{ij}, \ldots, R_{ij}}{\Lambda^{b_0}_{(1,3)}} \ , \tag{2.C.3}$$

where $b_0 = N - N_F + 4 = N + 1$ and the subscript $(1,3)$ denotes the matter content $(N_{\widetilde{A}}, N_F)$. The polynomial must of course respect the various symmetries of the theory. More significantly, the F-flatness equations following from $W^{\mathrm{dyn}}_{(1,3)}$ can be written in a Λ independent form. By setting $\Lambda = 0$, one sees that the equations must reduce to the classical constraints which follow upon expressing the gauge invariant field $\widetilde{O}_n$, M_{ij}, $\ldots$, R_{ij} in terms of their constituents Q_i and $\widetilde{A}$. Quantum corrections to these classical constraints are forbidden by symmetry

and holomorphy. These requirements fix $W^{\text{dyn}}_{(1,3)}$ up to an overall normalization. Once we have obtained $W^{\text{dyn}}_{(1,3)}$, we can derive the desired $W^{\text{dyn}}_{(1,0)}$ by giving mass to Q_i and integrating them out.

For $Sp(8)$, the above procedure uniquely determines the superpotential to be (this result appears in [22])

$$\begin{aligned} W^{\text{dyn}}_{(1,3)} = \frac{1}{\Lambda^9_{(1,3)}} \Big[& 1152(PPP) + 6912(RPN) + 3456(RRM) - 864\widetilde{O}_2(PNN) \\ & - 1728\widetilde{O}_2(RNM) + 108\widetilde{O}_2^2(NNM) - 108\widetilde{O}_2^2(PMM) + 9\widetilde{O}_2^3(MMM) \\ & + 192\widetilde{O}_3(NNN) - 576\widetilde{O}_3(RMM) + 144\widetilde{O}_2\widetilde{O}_3(NMM) + 32\widetilde{O}_3^2(MMM) \\ & + 432\widetilde{O}_4(NNM) + 432\widetilde{O}_4(PMM) - 36\widetilde{O}_2\widetilde{O}_4(MMM) \Big], \end{aligned} \tag{2.C.4}$$

up to normalization, where $(ABC) \equiv \epsilon^{ijklmn} A_{ij} B_{kl} C_{mn}$. Now that we have obtained $W^{\text{dyn}}_{(1,3)}$, we can integrate out Q_i by adding a mass term

$$W^{\text{dyn}}_{(1,3)} \to W^{\text{dyn}}_{(1,3)} + \frac{\mu^{ij}}{2} M_{ij}. \tag{2.C.5}$$

When solving the F-flatness condition, we can assume that $M_{ij}, N_{ij}, P_{ij}, R_{ij} \propto (\mu^{-1})_{ij}$ since μ^{ij} is the only quantity they can depend on. Plugging back in, we obtain the $Sp(8)$ superpotential (2.4.5) and (2.4.6). The same procedure leads to the superpotential (2.4.3) and (2.4.4) for $Sp(4)$ and $Sp(6)$, respectively [21, 22].

CHAPTER 3

The generalized Konishi anomaly approach

We derive the Konishi anomaly equations for $\mathcal{N} = 1$ supersymmetric gauge theories based on the classical gauge groups with matter in two-index tensor and fundamental representations, thus extending the existing results for $U(N)$. A general formula is obtained which expresses solutions to the Konishi anomaly equation in terms of solutions to the loop equations of the corresponding matrix model. This provides an alternative to the diagrammatic proof that the perturbative part of the glueball superpotential W_{eff} for these matter representations can be computed from matrix model integrals, and further shows that the two approaches always give the same result. The anomaly approach is found to be computationally more efficient in the cases we studied. Also, we show in the anomaly approach how theories with a traceless two-index tensor can be solved using an associated theory with a traceful tensor and appropriately chosen coupling constants.

3.1 Introduction

The recently established connection [5, 6, 7] between matrix models and the effective superpotentials of certain $\mathcal{N} = 1$ gauge theories provides us with a new tool for studying supersymmetric field theories. The connection, originally formulated in the context of $U(N)$ gauge theories with adjoint matter, has been established

following two distinct approaches, one based on superspace diagrammatics [13], and the other on generalized Konishi anomalies [17]. These derivations were subsequently generalized to a few more gauge groups and matter representations, but the list of examples is actually quite short at present. In particular, the diagrammatic approach has been applied to the classical gauge groups with matter in arbitrary two-index representations [14, 39, 40, 23, 46], while the anomaly approach has so far been used for $U(N)$ with matter in the adjoint and fundamental representations [17, 47], in (anti)symmetric tensor representations [46], and to quiver theories [46, 48]. [1] So basic questions remain regarding the general applicability of these ideas, and also whether matrix models can in fact successfully reproduce the known physics of supersymmetric gauge theories.

In [23], theories based on the classical gauge groups with two-index tensor matter were considered using the diagrammatic approach. In the case[2] of $Sp(N)$ with anti-symmetric matter, a comparison was made against an independently derived dynamical superpotential [21, 22] governing these theories. The comparison revealed agreement up to $h-1$ loops in perturbation theory (h is the dual Coxeter number), and a disagreement at h loops and beyond. Although it seemed most likely that the disagreement was due to nonperturbative effects, even at the perturbative level there were a number of subtleties deserving of further scrutiny. These subtleties mainly concern the class of diagrams which should be kept in the evaluation of the superpotential, and whether one is allowed to use Lie algebra identities to express objects of the form $\mathrm{Tr}(\mathcal{W}_\alpha)^{2h}$ in terms of lower traces including the glueball superfield $S \sim \mathrm{Tr}(\mathcal{W}_\alpha)^2$. Since these subtleties arise at the same order in perturbation theory as the observed discrepancies, it seems important to gain a better understanding of them. One motivation for the present

[1]The Konishi anomalies have also been applied without direct reference to a matrix model in [49].

[2]Our convention is such that $Sp(2) \cong SU(2)$.

work was to rederive the results of [23] in the anomaly approach to see if this gives the same result, and if so, to see which diagrams are effectively being computed. We will see that the anomaly approach corresponds to keeping at most two $\mathcal{W}_\alpha$'s per index loop and not using Lie algebra identities. So using these rules, whether one computes using diagrams or anomalies, one finds the same agreements/discrepancies between the gauge theory and the matrix model.

Another motivation for this work was to apply the anomaly approach to a wider class of theories. For the classical gauge groups with certain two-index tensors plus fundamentals, we will show how solutions to the Konishi anomaly equations can be obtained from solutions to the loop equations of the corresponding matrix model. This leads to the following general formula for the perturbative contribution to the effective glueball superpotential

$$W_{\text{eff}} = N\frac{\partial}{\partial S}\mathcal{F}_{S^2} + \frac{w_\alpha w^\alpha}{2}\frac{\partial^2}{\partial S^2}\mathcal{F}_{S^2} + 4\mathcal{F}_{\mathbb{RP}^2} + \mathcal{F}_{D^2} \tag{3.1.1}$$

where the $\mathcal{F}$'s are matrix model contributions of a given topology to the free energy. This formula generalizes the $U(N)$ results of [17, 47, 46], as well as results [14, 39, 40, 23, 46] found using the diagrammatic approach.

In fact, the above formula is only directly applicable to cases in which no tracelessness condition is imposed on the two-index tensors. In [23] it was shown that imposing a tracelessness condition requires one to include additional disconnected matrix model diagrams, and there was no simple formula relating the superpotential to the free energy of the traceless matrix model. On the other hand, one expects that the traceful theory should contain all the information about the traceless case provided one includes a Lagrange multiplier field to set the trace to zero. We will show how this works in detail, and find that indeed, the superpotential of the traceless theory can be extracted from the free energy of the traceful matrix model. We use this to rederive and extend some results

from [23] in a much more convenient fashion.

The remainder of this paper is organized as follows. In sections 3.2 and 3.3 we derive the gauge theory Konishi anomaly equations and the matrix model loop equations for the theories of interest. The theories can all be treated in a uniform way by using appropriate projection operators. In section 3.4 we discuss some of the subtleties alluded to above, and then go on to show that solutions to the gauge theory anomaly equations follow from those of the matrix model loop equations. Section 3.5 concerns the effects of tracelessness. Details of some of our calculations are given in appendices 3.A and 3.B.

Note: As we were preparing the manuscript, [30] appeared which overlaps with some of our discussion.

3.2 Loop equations on the gauge theory side

In this section we derive the gauge theory loop equations for various gauge groups and matter representations, extending the $U(N)$ result of [17, 47].

3.2.1 Setup

We consider an $\mathcal{N} = 1$ supersymmetric gauge theory with tree level superpotential

$$W_{\text{tree}} = \text{Tr}[W(\Phi)] + \widetilde{Q}_{\tilde{f}} m_{\tilde{f}f}(\Phi) Q_f, \qquad (3.2.1)$$

where the two-index tensor Φ_{ij} is in one of the following representations:

- $U(N)$ adjoint.
- $SU(N)$ adjoint.

- $SO(N)$ antisymmetric tensor.
- $SO(N)$ symmetric tensor, traceful or traceless.
- $Sp(N)$ symmetric tensor.
- $Sp(N)$ antisymmetric tensor, traceful or traceless.

For other $U(N)$ representations, see [50, 46]. In the Sp cases, the object with the denoted symmetry is related to Φ by

$$\Phi = \begin{cases} SJ & S_{ij}\text{: symmetric tensor,} \\ AJ & A_{ij}\text{: antisymmetric tensor.} \end{cases} \tag{3.2.2}$$

Here J is the invariant antisymmetric tensor of $Sp(N)$, namely

$$J_{ij} = \begin{pmatrix} 0 & \mathbb{1}_{N/2} \\ -\mathbb{1}_{N/2} & 0 \end{pmatrix}. \tag{3.2.3}$$

The tracelessness of the Sp antisymmetric tensor is defined with respect to this J, i.e., by $\mathrm{Tr}[AJ] = 0$.

Also, Q_f and $\widetilde{Q}_f$ are fundamental matter fields, with f and $\widetilde{f}$ being flavor indices. In the $U(N)$ case we have N_f fundamentals Q_f and N_f anti-fundamentals $\widetilde{Q}_{\widetilde{f}}$, while in the SO/Sp case we have N_f fundamentals Q_f. In the SO/Sp case, $\widetilde{Q}_{\widetilde{f}}$ is not an independent field but related to Q_f by

$$(\widetilde{Q}_{\widetilde{f}})_i = \begin{cases} (Q_{\widetilde{f}})_i & SO(N), \\ (Q_{\widetilde{f}})_j J_{ji} & Sp(N). \end{cases} \tag{3.2.4}$$

In the Sp case, N_f should be taken to be even to avoid the Witten anomaly [51].

W and m are taken to be polynomials

$$W(z) = \sum_{p=1}^{n} \frac{g_p}{p} z^p, \qquad m_{\widetilde{f}f}(z) = \sum_{p=1}^{n'} \frac{(m_p)_{\widetilde{f}f}}{p} z^p, \tag{3.2.5}$$

where in the traceless cases the $p = 1$ term is absent from $W(z)$. Further, due to the symmetry properties of the matrix Φ, some g_p vanish for certain representations:

$$g_{2p+1} = 0 \quad (p = 0, 1, 2, \cdots) \qquad \text{for } SO \text{ antisymmetric / } Sp \text{ symmetric.} \tag{3.2.6}$$

The symmetry properties of Φ also imply that the matrices $(m_p)_{\tilde{f}f}$ have the following symmetry properties:

$$(m_p)_{f'f} = \begin{cases} (-1)^p (m_p)_{ff'} & SO \text{ antisymmetric,} \\ (m_p)_{ff'} & SO \text{ symmetric,} \\ (-1)^{p+1} (m_p)_{ff'} & Sp \text{ symmetric,} \\ -(m_p)_{ff'} & Sp \text{ antisymmetric.} \end{cases} \tag{3.2.7}$$

In this and the next few sections, we discuss traceful cases only, postponing the traceless cases to section 3.5 (we regard the $SU(N)$ case as the traceless $U(N)$ case).

3.2.2 The loop equations

We will be interested in expectation values of chiral operators. As in [17, 47],

$$\{\mathcal{W}_\alpha, \mathcal{W}_\beta\} = [\Phi, \mathcal{W}_\alpha] = \mathcal{W}_\alpha Q = \widetilde{Q} \mathcal{W}_\alpha = 0 \tag{3.2.8}$$

in the chiral ring. Therefore, the complete list of independent single-trace chiral operators are $\mathrm{Tr}[\Phi^p]$, $\mathrm{Tr}[\mathcal{W}_\alpha \Phi^p]$, $\mathrm{Tr}[\mathcal{W}^2 \Phi^p]$, and $\widetilde{Q}_{\tilde{f}} \Phi^p Q_f$. As is standard, we define

$$S = -\frac{1}{32\pi} \mathrm{Tr}[\mathcal{W}_\alpha \mathcal{W}^\alpha], \quad w_\alpha = \frac{1}{4\pi} \mathrm{Tr}[\mathcal{W}_\alpha]. \tag{3.2.9}$$

The chiral operators can be packaged concisely in terms of the resolvents

$$\begin{aligned} R(z) &\equiv -\frac{1}{32\pi^2}\left\langle \mathrm{Tr}\left[\frac{\mathcal{W}^2}{z-\Phi}\right]\right\rangle, \qquad & w_\alpha(z) &\equiv \frac{1}{4\pi}\left\langle \mathrm{Tr}\left[\frac{\mathcal{W}_\alpha}{z-\Phi}\right]\right\rangle, \\ T(z) &\equiv \left\langle \mathrm{Tr}\left[\frac{1}{z-\Phi}\right]\right\rangle, \qquad & M_{f\tilde{f}}(z) &\equiv \left\langle \widetilde{Q}_{\tilde{f}}\frac{1}{z-\Phi}Q_f\right\rangle . \end{aligned} \tag{3.2.10}$$

Note that the indices of $M_{f\tilde{f}}$ are reversed relative to $\widetilde{Q}_{\tilde{f}}$, Q_f. The resolvent $w_\alpha(z)$ is nonvanishing only for $U(N)$; in all other cases $w_\alpha(z) \equiv 0$. This can be understood as follows. In these semi-simple cases the Lie algebra generators are traceless, so we cannot have a nonzero background field w_α. There being no preferred spinor direction specified by the background w_α, the spinor $w_\alpha(z)$ can be nothing but zero. Alternatively, if we integrate out Φ, then $w_\alpha(z)$ should be of the form $\langle \mathrm{Tr}[\mathcal{W}_\alpha](\mathrm{Tr}[\mathcal{W}^2])^n\rangle$ by the chiral ring relations (3.2.8). If we use the factorization property of chiral operator expectation values, this is proportional to $w_\alpha S^n$, which vanishes.

The resolvents defined in equation (3.2.10) provide sufficient data to determine the effective superpotential up to a coupling independent part, because of the relation

$$\langle \mathrm{Tr}[\Phi^p]\rangle = p\frac{\partial}{\partial g_p}W_{\mathrm{eff}}, \qquad \left\langle \widetilde{Q}_{\tilde{f}}\Phi^p Q_f\right\rangle = p\frac{\partial}{\partial (m_p)_{\tilde{f}f}}W_{\mathrm{eff}}. \tag{3.2.11}$$

The generalized Konishi [20] anomaly equation [17, 47, 49] is obtained by considering the divergence of the current associated with the variation of a particular field Ψ_a:

$$\delta\Psi_a = f_a, \tag{3.2.12}$$

where a is a gauge index. Then the anomaly equation reads

$$\left\langle \frac{\partial W_{\mathrm{tree}}}{\partial \Psi_a} f_a\right\rangle + \frac{1}{32\pi^2}\left\langle [\mathcal{W}_\alpha\mathcal{W}^\alpha]_a^b\frac{\partial f_b}{\partial\Psi_a}\right\rangle = 0, \tag{3.2.13}$$

where $\mathcal{W}_\alpha$ is in the representation furnished by Ψ. The first term in (3.2.13) represents the classical change of the action under the variation (3.2.12), while the second term in (3.2.13) corresponds to the quantum variation due to the change in the functional measure.

In the $U(N)$ case considered in [17, 47], there is no additional symmetry imposed on the field Φ, so $\delta\Phi_{ij} = f_{ij}$ can be any function of $\mathcal{W}_\alpha$ and Φ. In general, the tensor Φ will have some symmetry properties (symmetric or antisymmetric tensor in the present SO/Sp study), and f_{ij} should be chosen to reflect those. Similarly, the derivative $\partial/\partial\Psi_a = \partial/\partial\Phi_{ij}$ should be defined in accord with the symmetry property of Φ_{ij}. To this end, we define a projector P appropriate to each case:

$$P_{ij,kl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \sigma t_{il}t_{jk}), \tag{3.2.14}$$

where

$$\begin{cases} t_{ij} = \delta_{ij},\ \sigma = -1 & SO \text{ antisymmetric}, \\ t_{ij} = \delta_{ij},\ \sigma = +1 & SO \text{ symmetric}, \\ t_{ij} = J_{ij},\ \sigma = -1 & Sp \text{ symmetric}, \\ t_{ij} = J_{ij},\ \sigma = +1 & Sp \text{ antisymmetric}\ . \end{cases} \tag{3.2.15}$$

The tensor Φ_{ij} satisfies $P_{ij,kl}\Phi_{kl} = \Phi_{ij}$. Then, the symmetry property of $\delta\Phi$ discussed above is implemented by the replacements

$$f_a = f_{ij} \to P_{ij,kl}f_{kl}, \qquad \frac{\partial}{\partial\Psi_a} = \frac{\partial}{\partial\Phi_{ij}} \to P_{ij,kl}\frac{\partial}{\partial\Phi_{kl}}. \tag{3.2.16}$$

With this replacement, f_{ij} can be any function of $\mathcal{W}_\alpha$ and Φ as in the $U(N)$ case. The derivative can be treated as in the $U(N)$ case also.

There is no such issue for the Q and $\widetilde{Q}$ fields, although we have to remember that they are not independent for SO/Sp.

With the projectors in hand, there is no difficulty in deriving the loop equations for SO/Sp. Here we just present the resulting loop equations, leaving the details to Appendix 3.A:

$$
\begin{aligned}
[W'R]_- &= \tfrac{1}{2}R^2, \\
[W'T + \mathrm{tr}(m'M)]_- &= \begin{cases} \left(T - \frac{2}{z}\right) R & SO \text{ antisymmetric}, \\ \left(T - 2\frac{d}{dz}\right) R & SO \text{ symmetric }, \\ \left(T + \frac{2}{z}\right) R & Sp \text{ symmetric }, \\ \left(T + 2\frac{d}{dz}\right) R & Sp \text{ antisymmetric}, \end{cases} \\
2[(Mm)_{ff'}]_- &= R\delta_{ff'}, \\
2[(mM)_{\widetilde{f}\widetilde{f}'}]_- &= R\delta_{\widetilde{f}\widetilde{f}'},
\end{aligned}
\tag{3.2.17}
$$

where $[F(z)]_-$ means to drop non-negative powers in a Laurent expansion in z. The last two equations are really the same equation due to the symmetry properties of m (see equation (3.2.7)), and Φ. Note that there is no $w_\alpha(z)$ in these cases as explained below Eq. (3.2.10). For the sake of comparison, the $U(N)$ loop equations are [17, 47]

$$
\begin{aligned}
[W'R]_- &= R^2, \\
[W'w_\alpha]_- &= 2w_\alpha R, \\
[W'T + \mathrm{tr}(m'M)]_- &= 2TR + w_\alpha w^\alpha, \\
[(Mm)_{ff'}]_- &= R\delta_{ff'}, \\
[(mM)_{\widetilde{f}\widetilde{f}'}]_- &= R\delta_{\widetilde{f}\widetilde{f}'}.
\end{aligned}
\tag{3.2.18}
$$

One observes some extra numerical factors in the SO/Sp case as compared to the $U(N)$ case. The $\frac{1}{2}$ in the first equation is from the $\frac{1}{2}$ in the definition of $P_{ij,kl}$, while the factor 2 in the last two equations is because in the SO/Sp case Q and $\widetilde{Q}$ are really the same field, so the variation of $\widetilde{Q}mQ$ under δQ for SO/Sp is twice

as large as that for $U(N)$. Finally, the $\frac{1}{z}R(z)$ and $\frac{d}{dz}R(z)$ terms in the second equation of (3.2.17) come from the second term of $P_{ij,kl}$.

The solution to the loop equations (3.2.17) or (3.2.18) is determined uniquely [17] given the condition

$$S = \oint_C \frac{dz}{2\pi i} R(z), \qquad w_\alpha = \oint_C \frac{dz}{2\pi i} w_\alpha(z), \qquad N = \oint_C \frac{dz}{2\pi i} T(z), \tag{3.2.19}$$

where the second equation is only for the $U(N)$ case. The contour C goes around the critical point of $W(z)$. Therefore, if we recall the relation (3.2.11), we can say that the loop equations are all we need to determine the superpotential W_{eff}.

3.3 Loop equations on the matrix model side

Let us consider the matrix model which corresponds to the gauge theory in the previous section. Its partition function is

$$\mathbf{Z} = e^{-\frac{1}{\mathbf{g}^2}\mathbf{F}(\mathbf{S})} = \int d\mathbf{\Phi} d\mathbf{Q} d\tilde{\mathbf{Q}}\, e^{-\frac{1}{\mathbf{g}} W_{\text{tree}}(\mathbf{\Phi}, \mathbf{Q}, \tilde{\mathbf{Q}})}. \tag{3.3.1}$$

We denote matrix model quantities by boldface letters. Here, $\mathbf{\Phi}$ is an $\mathbf{N} \times \mathbf{N}$ matrix with the same symmetry property as the corresponding matter field in the gauge theory. $\mathbf{Q}_f$ and $\tilde{\mathbf{Q}}_{\tilde{f}}$ are defined in a similar way to their gauge theory counterparts (therefore $d\tilde{\mathbf{Q}}$ in (3.3.1) is not included for SO/Sp). The function (or the "action") W_{tree} is the one defined in (3.2.1). We will take the $\mathbf{N} \to \infty$, $\mathbf{g} \to 0$ limit with the 't Hooft coupling $\mathbf{S} = \mathbf{g}\mathbf{N}$ kept fixed. The dependence of the free energy $\mathbf{F}(\mathbf{S})$ on $\mathbf{N}$ is eliminated using the relation $\mathbf{N} = \mathbf{S}/\mathbf{g}$, and we expand $\mathbf{F}(\mathbf{S})$ as

$$\mathbf{F}(\mathbf{S}) = \sum_{\mathcal{M}} \mathbf{g}^{2-\chi(\mathcal{M})} \mathcal{F}_{\mathcal{M}}(\mathbf{S}) = \mathcal{F}_{S^2} + \mathbf{g}\mathcal{F}_{\mathbb{RP}^2} + \mathbf{g}\mathcal{F}_{D^2} + \cdots, \tag{3.3.2}$$

where the sum is over all compact topologies $\mathcal{M}$ of the matrix model diagrams written in the 't Hooft double-line notation, and $\chi(\mathcal{M})$ is the Euler number of

$\mathcal{M}$. The cases which will be of interest to us are the sphere S^2, projective plane $\mathbb{RP}^2$, and disk D^2, with $\chi = 2, 1$, and 1, respectively. All other contributions have $\chi \leq 0$.

We define matrix model resolvents as follows:

$$\mathbf{R}(z) \equiv \mathbf{g}\left\langle \mathrm{Tr}\left[\frac{1}{z-\mathbf{\Phi}}\right]\right\rangle, \qquad \mathbf{M}_{f\tilde{f}}(z) \equiv \mathbf{g}\left\langle \tilde{\mathbf{Q}}_{\tilde{f}} \frac{1}{z-\mathbf{\Phi}} \mathbf{Q}_f \right\rangle. \tag{3.3.3}$$

These resolvents provide sufficient data to determine the free energy $\mathbf{F}$ up to a coupling independent part since

$$\mathbf{g}\langle \mathrm{Tr}[\mathbf{\Phi}^p]\rangle = p\frac{\partial}{\partial g_p}\mathbf{F}, \qquad \mathbf{g}\left\langle \tilde{\mathbf{Q}}_{\tilde{f}} \mathbf{\Phi}^p \mathbf{Q}_f \right\rangle = p\frac{\partial}{\partial (m_p)_{\tilde{f}f}}\mathbf{F}. \tag{3.3.4}$$

We expand the resolvents in topologies just as we did for $\mathbf{F}$:

$$\mathbf{R}(z) = \sum_{\mathcal{M}} \mathbf{g}^{2-\chi(\mathcal{M})}\mathbf{R}_{\mathcal{M}}(z), \qquad \mathbf{M}(z) = \sum_{\mathcal{M}} \mathbf{g}^{2-\chi(\mathcal{M})}\mathbf{M}_{\mathcal{M}}(z). \tag{3.3.5}$$

Although $\mathbf{R}_{\mathbb{RP}^2}$ and $\mathbf{R}_{D^2}$ are of the same order in $\mathbf{g}$, they can be distinguished unambiguously because all terms in $\mathbf{R}_{D^2}$ contains coupling constants $m_{\tilde{f}f}$, while $\mathbf{R}_{\mathbb{RP}^2}$ does not depend on them at all. This is easily seen in the diagrammatic expansion of $\mathbf{F}$. Also, because $\mathbf{F}_{S^2}$ and $\mathbf{F}_{\mathbb{RP}^2}$ do not contain m, the expansion of $\mathbf{M}$ starts from the disk contribution, $\mathbf{M}_{D^2}$.

Now we can derive the matrix model loop equations. Consider changing the integration variables as

$$\delta\mathbf{\Psi}_a = \mathbf{f}_a. \tag{3.3.6}$$

Since the partition function is invariant under this variation, we obtain

$$0 = -\frac{1}{\mathbf{g}}\frac{\partial W_{\mathrm{tree}}}{\partial \mathbf{\Psi}_a}\mathbf{f}_a + \frac{\partial \mathbf{f}_a}{\partial \mathbf{\Psi}_a}. \tag{3.3.7}$$

The first term came from the change in the "action" and corresponds to the first term (the classical variation) of the generalized Konishi anomaly equation

(3.2.13). On the other hand, the second term came from the Jacobian and corresponds to the second term (the anomalous variation) of Eq. (3.2.13).

The derivation of the loop equations now can be done exactly in parallel to the derivation of the gauge theory loop equations. In the SO/Sp case, we again have to consider the projector $P_{ij,kl}$. Here we leave details of the derivation to Appendix 3.B and present the results. For SO/Sp, they are

$$\begin{aligned}
&\mathbf{g}\left\langle \mathrm{Tr}\frac{W'(\mathbf{\Phi})}{z-\mathbf{\Phi}}\right\rangle + \mathbf{g}\left\langle \tilde{\mathbf{Q}}\frac{m'(\mathbf{\Phi})}{z-\mathbf{\Phi}}\mathbf{Q}\right\rangle \\
&\qquad\qquad = \frac{1}{2}\left\langle\left(\mathbf{g}\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right)^2\right\rangle \pm \frac{\sigma}{2}\mathbf{g}^2\left\langle \mathrm{Tr}\frac{1}{(z-\mathbf{\Phi})(z-\sigma\mathbf{\Phi})}\right\rangle, \\
&2\left\langle \tilde{\mathbf{Q}}_{\tilde{f}}\frac{m_{\tilde{f}f}(\mathbf{\Phi})}{z-\mathbf{\Phi}}\mathbf{Q}_{f'}\right\rangle = \mathbf{g}\left\langle \mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right\rangle \delta_{ff'}, \\
&2\left\langle \tilde{\mathbf{Q}}_{\tilde{f}}\frac{m_{\tilde{f}'f}(\mathbf{\Phi})}{z-\mathbf{\Phi}}\mathbf{Q}_{f}\right\rangle = \mathbf{g}\left\langle \mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right\rangle \delta_{\tilde{f}\tilde{f}'},
\end{aligned} \tag{3.3.8}$$

in the SO and Sp cases, respectively. The last two equations are really the same because of the symmetry properties of $\mathbf{\Phi}$ and $m_{\tilde{f}f}$.

Equations (3.3.8) include terms of all orders in $\mathbf{g}$. Expanding the matrix model expectation values in powers of $\mathbf{g}$, plugging in the expansion (3.3.5) and comparing the $\mathcal{O}(1)$ and $\mathcal{O}(\mathbf{g}^1)$ terms, we obtain the SO/Sp loop equations[3].

[3] In the SO antisymmetric and Sp symmetric cases, $\mathbf{R}_{\mathbb{RP}^2}$ can be expressed [40, 39] in terms of $\mathbf{R}_{S^2}$, which leads to the expression

$$\mathcal{F}_{S^2}(\mathbf{S}) = \mp\frac{1}{2}\frac{\partial}{\partial \mathbf{S}}\mathcal{F}_{\mathbb{RP}^2} \tag{3.3.9}$$

in the SO and Sp cases, respectively.

This is done in Appendix 3.B, and the results are:

$$
\begin{aligned}
[W'\mathbf{R}_{S^2}]_- &= \tfrac{1}{2}(\mathbf{R}_{S^2})^2 \\
[W'\mathbf{R}_{\mathbb{RP}^2}]_- &= \begin{cases} \left(\mathbf{R}_{\mathbb{RP}^2} - \frac{1}{2z}\right)\mathbf{R}_{S^2} & SO \text{ antisymmetric} \\ \left(\mathbf{R}_{\mathbb{RP}^2} - \frac{1}{2}\frac{d}{dz}\right)\mathbf{R}_{S^2} & SO \text{ symmetric} \\ \left(\mathbf{R}_{\mathbb{RP}^2} + \frac{1}{2z}\right)\mathbf{R}_{S^2} & Sp \text{ symmetric} \\ \left(\mathbf{R}_{\mathbb{RP}^2} + \frac{1}{2}\frac{d}{dz}\right)\mathbf{R}_{S^2} & Sp \text{ antisymmetric} \end{cases} \\
[W'\mathbf{R}_{D^2} + \mathrm{tr}(m'\mathbf{M}_{D^2})]_- &= \mathbf{R}_{D^2}\mathbf{R}_{S^2} \\
2[(\mathbf{M}_{D^2}m)_{ff'}]_- &= \mathbf{R}_{S^2}\delta_{ff'} \\
2[(m\mathbf{M}_{D^2})_{\tilde{f}\tilde{f}'}]_- &= \mathbf{R}_{S^2}\delta_{\tilde{f}\tilde{f}'},
\end{aligned}
\tag{3.3.10}
$$

We separated the $\mathbf{R}_{\mathbb{RP}^2}$ and $\mathbf{R}_{D^2}$ contributions using the difference in their dependence on $m_{\tilde{f}f}$ (see the argument below Eq. (3.3.5)). Again, the last two equations are really the same equation. For comparison, the $U(N)$ loop equations are

$$
\begin{aligned}
[W'\mathbf{R}_{S^2}]_- &= (\mathbf{R}_{S^2})^2 \\
[W'\mathbf{R}_{D^2} + \mathrm{tr}(m'\mathbf{M}_{D^2})]_- &= 2\mathbf{R}_{D^2}\mathbf{R}_{S^2} \\
[(\mathbf{M}_{D^2}m)_{ff'}]_- &= \mathbf{R}_{S^2}\delta_{ff'} \\
[(m\mathbf{M}_{D^2})_{\tilde{f}\tilde{f}'}]_- &= \mathbf{R}_{S^2}\delta_{\tilde{f}\tilde{f}'},
\end{aligned}
\tag{3.3.11}
$$

Note that there is no $\mathbb{RP}^2$ contribution for $U(N)$.

The solutions to equations (3.3.10) or (3.3.11) are determined uniquely given the condition

$$
\mathbf{S} = \oint_C \frac{dz}{2\pi i}\mathbf{R}_{S^2}(z), \qquad 0 = \oint_C \frac{dz}{2\pi i}\mathbf{R}_{\mathbb{RP}^2}(z), \qquad 0 = \oint_C \frac{dz}{2\pi i}\mathbf{R}_{D^2}(z). \tag{3.3.12}
$$

In this sense, the loop equations are all we need to determine the free energy $\mathbf{F}$.

3.4 Connection between gauge theory and matrix model resolvents

On the gauge theory side we have arrived at the loop equations (3.2.17). If we can solve these equations for the resolvents, in particular for $T(z)$, we will have sufficient data to determine the glueball superpotential $W_{\mathrm{eff}}(S)$ up to a coupling independent part. In [17], it was shown for $U(N)$ with adjoint matter that the solution can be obtained with the help of an auxiliary matrix model. On the other hand, in [13, 14, 23, 46] it was proved by perturbative diagram expansion that, for $U(N)$ and SO/Sp with two-index tensor matter, if one only inserts up to two field strength superfield $\mathcal{W}_\alpha$'s per index loop then the calculation of $W_{\mathrm{eff}}(S)$ reduces to matrix integrals.

However, there are a number of reasons to study further the relation between the gauge theory and matrix model loop equations. First, as pointed out in [23] (see also p.11 of [17], and [44]), there are subtleties in using chiral ring relations at order S^h and higher, where h is the dual Coxeter number of the gauge group, and these could be related to the discrepancies observed in [23]. Since traces of schematic form $\mathrm{Tr}[(\mathcal{W}_\alpha^2)^n]$ $(n \geq h)$ can be rewritten in terms of lower power traces at these orders, imposing chiral ring relations *before* using the equation of motion of S is not necessarily justified. So, it is important to clarify how this subtlety is treated in the Konishi anomaly approach. Second, as a practical matter, the anomaly approach is more efficient than the diagrammatic approach in the cases we studied.

So, let us adopt the following point of view (some related ideas were explored in [49]). Let us not assume the reduction to a matrix model *a priori*. Then the gauge theory resolvents R, T, and M are just unknown functions that enable us

to determine the coupling dependent part of the glueball effective action. We do know that we can evaluate the perturbative contribution to them by Feynman diagrams, but we do not know whether they are affected by nonperturbative effects or whether they can be calculated using a matrix model. These resolvents satisfy the loop equations (3.2.17), and given the conditions (3.2.19), they are determined uniquely. Similarly, the matrix model resolvents $\mathbf{R}_{S^2}$, $\mathbf{R}_{\mathbb{RP}^2}$, $\mathbf{R}_{D^2}$, and $\mathbf{M}_{D^2}$ are now just functions satisfying matrix model loop equations (3.3.10). If we impose the condition (3.3.12), these resolvents are also determined uniquely, and by definition can be evaluated in matrix model perturbation theory.

Now, let us ask what the relation between the two sets of resolvents is. Actually it is simple: if we know the matrix model resolvents, we can construct the gauge theory resolvents as follows. In the SO/Sp case,

$$\begin{aligned} R(z) &= \mathbf{R}_{S^2}(z), \\ T(z) &= N\frac{\partial}{\partial \mathbf{S}}\mathbf{R}_{S^2}(z) + 4\mathbf{R}_{\mathbb{RP}^2}(z) + \mathbf{R}_{D^2}(z), \\ M(z) &= \mathbf{M}_{D^2}(z) \end{aligned} \tag{3.4.1}$$

with S and $\mathbf{S}$ identified; in the $U(N)$ case, we get

$$\begin{aligned} R(z) &= \mathbf{R}_{S^2}(z), \qquad w_\alpha(z) = w_\alpha\frac{\partial}{\partial \mathbf{S}}\mathbf{R}_{S^2}(z), \\ T(z) &= N\frac{\partial}{\partial \mathbf{S}}\mathbf{R}_{S^2}(z) + \frac{w_\alpha w^\alpha}{2}\frac{\partial^2}{\partial \mathbf{S}^2}\mathbf{R}_{S^2}(z) + \mathbf{R}_{D^2}(z), \\ M(z) &= \mathbf{M}_{D^2}(z) \end{aligned} \tag{3.4.2}$$

with the same $S = \mathbf{S}$ identification.[4] One can easily check that if the matrix model resolvents satisfy the matrix model loop equations (3.3.10) or (3.3.11), then the gauge theory resolvents satisfy the gauge theory loop equations (3.2.17) or (3.2.18). The requirement (3.2.19) is also satisfied provided that the matrix

[4]Some of these relations have been written down in [8, 52, 53].

model resolvents satisfy the requirement (3.3.12). Further, these relations lead to

$$\begin{aligned}\langle \mathrm{Tr}[\Phi^p]\rangle_{\text{gauge theory}} &= p\frac{\partial}{\partial g_p}W_{\text{eff}} \\ &= p\frac{\partial}{\partial g_p}\left[N\frac{\partial}{\partial S}\mathcal{F}_{S^2} + \frac{w_\alpha w^\alpha}{2}\frac{\partial^2}{\partial S^2}\mathcal{F}_{S^2} + 4\mathcal{F}_{\mathbb{RP}^2} + \mathcal{F}_{D^2}\right], \quad (3.4.3)\end{aligned}$$

which implies a relation between the effective superpotential and the matrix model quantities:

$$W_{\text{eff}} = N\frac{\partial}{\partial S}\mathcal{F}_{S^2} + \frac{w_\alpha w^\alpha}{2}\frac{\partial^2}{\partial S^2}\mathcal{F}_{S^2} + 4\mathcal{F}_{\mathbb{RP}^2} + \mathcal{F}_{D^2} \quad (3.4.4)$$

up to a coupling independent additive part. This proves that the gauge theory diagrams considered in the Konishi anomaly approach reduce to matrix model integrals for all matter representations considered. Further, we do not have to take into account nonperturbative effects, since we can assume a perturbative expansion in the matrix model (although, strictly speaking, one should also verify that the Konishi anomalies receive no nonperturbative corrections [49]).

The relations (3.4.1) and (3.4.2) are consistent with inserting at most two $\mathcal{W}_\alpha$'s per index loop, but not with inserting more than two and then using Lie algebra relations. For instance, this can be seen from the diagrammatic expansion of $\mathbf{R}_{S^2}(z)$. So this shows us explicitly which diagrams are being computed in the Konishi anomaly approach.

In the $U(N)$ case [17], it was convenient to collect all the gauge theory resolvents into a "superfield" $\mathcal{R}$, because of the "supersymmetry" under a shift of $\mathcal{W}_\alpha$ by a Grassmann number, and one could relate $\mathcal{R}$ to the matrix model resolvent $\mathbf{R}_{S^2}$. This fact enabled one to extract all the gauge theory resolvents solely from $\mathbf{R}_{S^2}$. However, in more general cases this trick does not work, and we have to relate the two sets of resolvents directly as in (3.4.1).

3.5 Traceless cases

So far, we considered two-index traceful matter Φ_{ij}, and discussed the relation between the gauge theory and the corresponding matrix model. In this section, we consider traceless[5] tensors $\widetilde{\Phi}_{ij}$. These traceless tensors were studied in [23], and a method of evaluating the glueball effective superpotential $\widetilde{W}_{\rm eff}(S)$ from the combinatorics of the matrix model diagrams was given. However, the precise connection between the gauge theory and the matrix model quantities was not transparent, since one had to keep some of the matrix model diagrams and drop others in a way that seemed rather arbitrary from the matrix model point of view. Instead, here we show that the calculation of $\widetilde{W}_{\rm eff}(S)$ in gauge theory with *traceless* matter reduces to a *traceful* matrix model[6].

3.5.1 Traceless gauge theory vs. traceful matrix model

To derive the generalized Konishi anomaly equation for a traceless tensor we have to use the appropriate projector

$$\widetilde{P}_{ij,kl} \equiv P_{ij,kl} - \frac{1}{N}\delta_{ij}P_{mm,kl} = P_{ij,kl} - \frac{1}{N}\delta_{ij}\delta_{kl}, \tag{3.5.1}$$

where P is the projector of the corresponding traceful theory; the second equality holds for any projector defined in (3.2.14). The anomaly term (the second term of Eq. (3.2.13)) is the same as in the traceful case, since the trace part is a singlet and does not couple to the gauge field. Therefore, the only difference in the anomaly equation between traceful and traceless cases is in the classical

[5] In this section, we denote traceless quantities by tildes to distinguish them from their traceful counterparts.

[6] A connection between traceful and traceless gauge theories in the Leigh–Strassler deformed $\mathcal{N}=4$ $SU(N)$ theory was discussed in [54].

variation (the first term of Eq. (3.2.13)), namely

$$\mathrm{Tr}[(Pf)\widetilde{W}'(\Phi)] \to \mathrm{Tr}[(\widetilde{P}f)\widetilde{W}'(\Phi)] = \mathrm{Tr}[(Pf)\widetilde{W}'(\Phi)] - \frac{1}{N}\mathrm{Tr}[f]\mathrm{Tr}[\widetilde{W}'(\Phi)]. \tag{3.5.2}$$

For definiteness, let us focus on $SU(N)$ adjoint matter, which can be thought of as traceless $U(N)$ adjoint matter, without fundamentals added; we will generalize the discussion to other groups and matter representations afterward. In this case, the last term of Eq. (3.5.2) changes the $U(N)$ loop equation (the first and the third lines of (3.2.18)) to

$$[\widetilde{W}'(z)\widetilde{R}(z)]_- + g_1\widetilde{R}(z) = \widetilde{R}(z)^2, \qquad [\widetilde{W}'(z)\widetilde{T}(z)]_- + g_1\widetilde{T}(z) = 2\widetilde{R}(z)\widetilde{T}(z). \tag{3.5.3}$$

Note that $w_\alpha(z) = 0$ for $SU(N)$. The constant g_1 is

$$g_1 \equiv -\frac{1}{N}\Big\langle \mathrm{Tr}[\widetilde{W}'(\widetilde{\Phi})]\Big\rangle. \tag{3.5.4}$$

If we define

$$W(z) \equiv \widetilde{W}(z) + g_1 z, \tag{3.5.5}$$

the above equations are

$$[W'(z)\widetilde{R}(z)]_- = \widetilde{R}(z)^2, \qquad [W'(z)\widetilde{T}(z)]_- = 2\widetilde{R}(z)\widetilde{T}(z). \tag{3.5.6}$$

These are of the same form as the loop equations with *traceful* matter and the tree level superpotential W. Therefore, in order to obtain the effective glueball superpotential $\widetilde{W}_{\mathrm{eff}}(S)$ for traceless matter, we can instead solve the traceful theory with the shifted tree level superpotential W, choosing the value of g_1 appropriately. The solution to these loop equations is determined uniquely given the condition

$$S = \oint_C \frac{dz}{2\pi i}\widetilde{R}(z), \qquad N = \oint_C \frac{dz}{2\pi i}\widetilde{T}(z). \tag{3.5.7}$$

In the case of traceful matter, the contour is around a critical point of the tree level superpotential. However, for traceless matter, the loop equations above tell us that the contour should be taken around the critical point of the shifted superpotential (3.5.5), rather than the original $\widetilde{W}$. This is because we cannot change all the eigenvalues of $\widetilde{\Phi}$ independently due to the tracelessness condition $\mathrm{Tr}[\widetilde{\Phi}] = 0$.

Let the resolvents of the traceful theory with tree level superpotential $W(\Phi)$ be R and T, with g_1 treated as an independent variable. R and T are functions of z, $g_{p\geq 1}$ as well as S, N: $R = R(z; g_{p\geq 1}, S)$, $T = T(z; g_{p\geq 1}, S, N)$. We will often omit S and N in the arguments henceforth to avoid clutter. Since R and T satisfy the same loop equations as $\widetilde{R}$ and $\widetilde{T}$ provided g_1 is chosen appropriately, i.e. $g_1 = g_1(g_{p\geq 2}, S, N) \equiv \widetilde{g}_1$, it should be that

$$\widetilde{R}(z; g_{p\geq 2}) = R(z; g_{p\geq 1})\big|_{g_1=\widetilde{g}_1}, \qquad \widetilde{T}(z; g_{p\geq 2}) = T(z; g_{p\geq 1})\big|_{g_1=\widetilde{g}_1}. \tag{3.5.8}$$

These satisfy the conditions (3.5.7) given that R and T satisfy the conditions (3.5.7) without tildes. Expanding these in z, we find

$$\begin{aligned} \langle \mathrm{Tr}[\mathcal{W}^2 \widetilde{\Phi}^p]\rangle^{\text{traceless}}_{g_{p\geq 2}} &= \langle \mathrm{Tr}[\mathcal{W}^2 \Phi^p]\rangle^{\text{traceful}}_{g_{p\geq 1}}\big|_{g_1=\widetilde{g}_1}, \\ \langle \mathrm{Tr}[\widetilde{\Phi}^p]\rangle^{\text{traceless}}_{g_{p\geq 2}} &= \langle \mathrm{Tr}[\Phi^p]\rangle^{\text{traceful}}_{g_{p\geq 1}}\big|_{g_1=\widetilde{g}_1}. \end{aligned} \tag{3.5.9}$$

In particular, setting $p = 1$ in the second equation,

$$\langle \mathrm{Tr}[\Phi]\rangle^{\text{traceful}}_{g_{p\geq 1}}\big|_{g_1=\widetilde{g}_1} = \left[\frac{\partial}{\partial g_1} T(z; g_{p\geq 1})\right]\Bigg|_{g_1=\widetilde{g}_1} = 0, \tag{3.5.10}$$

which can be used for determining g_1 in terms of all other parameters.[7] We infer from Eq. (3.5.8) equality between the traceless and traceful effective superpotentials:

$$\widetilde{W}_{\text{eff}}(g_{p\geq 2}, S, N) = W_{\text{eff}}(g_{p\geq 1}, S, N)|_{g_1=\widetilde{g}_1(g_{p\geq 2}, S, N)}. \tag{3.5.11}$$

[7]One might have expected that g_1 can be determined by Eq. (3.5.4). However, it is easy to show using the relation (3.5.9) that the equation is just the equation of motion of the traceful theory, which is identically satisfied for any g_1: $0 \equiv \langle \mathrm{Tr}[W'(\Phi)]\rangle = \langle \mathrm{Tr}[\widetilde{W'}(\widetilde{\Phi})]\rangle + N g_1$.

As long as we impose the tracelessness condition (3.5.10), this correctly reproduces the relation (3.5.9). Note that $\widetilde{g}_1$ depends on N; this is the origin of the complicated N dependence of $\widetilde{W}_{\text{eff}}$ found in [23].

Because we know that the traceful theory can be solved by the associated traceful matrix model, we can calculate the effective superpotential using that matrix model. Specifically, in the present case, it is given in terms of the free energy of the traceful matrix model by

$$\widetilde{W}_{\text{eff}}(g_{p\geq 2}, S, N) = \left[N \frac{\partial}{\partial S} \mathcal{F}_{S^2} \right] \Bigg|_{g_1 = \widetilde{g}_1} . \tag{3.5.12}$$

The function $\widetilde{g}_1(g_2, g_3, \cdots, S, N)$ is determined by

$$\langle \mathrm{Tr}[\widetilde{\Phi}] \rangle = \left[N \frac{\partial}{\partial S} \frac{\partial}{\partial g_1} \mathcal{F}_{S^2} \right] \Bigg|_{g_1 = \widetilde{g}_1} = 0. \tag{3.5.13}$$

If we add fundamental fields, the shift constant g_1 is changed to

$$g_1 \equiv -\frac{1}{N} \left\langle \mathrm{Tr}[\widetilde{W}'(\widetilde{\Phi})] \right\rangle - \frac{1}{N} \left\langle \widetilde{Q}_{\tilde{f}} m_{\tilde{f}f} Q_f \right\rangle, \tag{3.5.14}$$

but everything else remains the same; we just have to work with the traceful theory and the shifted tree level superpotential. g_1 is determined by the tracelessness condition.

We only discussed the $SU(N)$ case in the above, but the generalization to other tensors, i.e., SO traceless symmetric tensor and Sp traceless antisymmetric tensor, is straightforward. We just shift the tree level superpotential as (3.5.5), and work with the traceful theory instead.

3.5.2 Examples

Here we explicitly demonstrate how the method outlined above works in the case of a cubic tree level superpotential,

$$\widetilde{W}(\widetilde{\Phi}) = \frac{m}{2} \widetilde{\Phi}^2 + \frac{g}{3} \widetilde{\Phi}^3. \tag{3.5.15}$$

The associated traceful tree level superpotential is

$$W(\Phi) = \lambda\Phi + \frac{m}{2}\Phi^2 + \frac{g}{3}\Phi^3 \tag{3.5.16}$$

$(g_1 = \lambda, g_2 = m, g_3 = g)$.

3.5.2.1 $SU(N)$ adjoint

We first consider $SU(N)$ with adjoint matter and no fundamentals. In [23] it was found by perturbative computation to order g^6 that the corresponding W_{eff} vanishes due to a cancellation among diagrams. We will now prove that $W_{\text{eff}} = 0$ to all orders in g.

The planar contributions to the free energy of the traceful matrix model can be computed exactly by the standard method [45]:

$$\mathcal{F}_{S^2} = SW_0 + \frac{1}{2}S^2 \ln\left(\frac{\tilde{m}}{\sqrt{1+y}m}\right) - \frac{2}{3}\frac{S^2}{y}\left[1 + \frac{3}{2}y + \frac{1}{8}y^2 - (1+y)^{3/2}\right] \tag{3.5.17}$$

with

$$\begin{aligned} \tilde{m} &= \sqrt{m^2 - 4\lambda g} \\ W_0 &= \frac{1}{2g}(\tilde{m} - m)\left(\lambda + \frac{1}{12g}(\tilde{m} - m)(\tilde{m} + 2m)\right) \\ \frac{y}{(1+y)^{3/2}} &= \frac{8g^2 S}{m^3}. \end{aligned} \tag{3.5.18}$$

We discarded some g independent contributions. The W_0 term arises from shifting Φ to eliminate the linear term in $W(\Phi)$. The superpotential is therefore

$$W_{\text{eff}} = N\frac{\partial \mathcal{F}_{S^2}}{\partial S} = NW_0 + \frac{NS}{6y}\left[-4 - 6y + 6y\ln\left(\frac{\tilde{m}}{m\sqrt{1+y}}\right) + 4(1+y)^{3/2}\right]. \tag{3.5.19}$$

Imposing $\partial W_{\text{eff}}/\partial\lambda = 0$ leads to, after some algebra,

$$\lambda = -\frac{2gS}{m}, \quad y = \frac{8g^2 S}{m^3}. \tag{3.5.20}$$

Substituting back into (3.5.19) and doing some more algebra, we find

$$W_{\text{eff}} = 0. \tag{3.5.21}$$

This vanishing of the perturbative contribution to the effective superpotential is consistent with the gauge theory analysis of [55]. In fact, it is shown there that $W_{\text{eff}} = 0$ for any tree level superpotential with only odd power interactions.

3.5.2.2 $Sp(N)$ antisymmetric tensor

Now consider $Sp(N)$ with an antisymmetric tensor and no fundamentals. By diagram calculations or by computer, the planar and $\mathbb{RP}^2$ contributions to the free energy of the traceful matrix model are

$$\begin{aligned}
\mathcal{F}_{S^2} = &-\frac{\lambda^2 S}{2m} - \left(\frac{\lambda S^2}{2m^2} + \frac{\lambda^3 S}{3m^3}\right) g - \left(\frac{S^3}{6m^3} + \frac{\lambda^2 S^2}{m^4} + \frac{\lambda^4 S}{2m^3}\right) g^2 \\
&- \left(\frac{\lambda S^3}{m^5} + \frac{8\lambda^3 S^2}{3m^6} + \frac{\lambda^5 S}{m^7}\right) g^3 - \left(\frac{S^4}{3m^6} + \frac{5\lambda^2 S^3}{m^7} + \frac{8\lambda^4 S^2}{m^8} + \frac{7\lambda^6 S}{3m^9}\right) g^4 \\
&- \left(\frac{4\lambda S^4}{m^8} + \frac{70\lambda^3 S^3}{3m^9} + \frac{128\lambda^5 S^2}{5m^{10}} + \frac{6\lambda^7 S}{m^{11}}\right) g^5 \\
&- \left(\frac{7S^5}{6m^9} + \frac{32\lambda^2 S^4}{m^{10}} + \frac{105\lambda^4 S^3}{m^{11}} + \frac{256\lambda^6 S^2}{3m^{12}} + \frac{33\lambda^8 S}{2m^{13}}\right) g^6 \\
&- \left(\frac{21\lambda S^5}{m^{11}} + \frac{640\lambda^3 S^4}{3m^{12}} + \frac{462\lambda^5 S^3}{m^{13}} + \frac{2048\lambda^7 S^2}{7m^{14}} + \frac{143\lambda^9 S}{3m^{15}}\right) g^7 \\
&- \left(\frac{16S^6}{3m^{12}} + \frac{231\lambda^2 S^5}{m^{13}} + \frac{1280\lambda^4 S^4}{m^{14}}\right. \\
&\qquad\qquad \left. + \frac{2002\lambda^6 S^3}{m^{15}} + \frac{1024\lambda^8 S^2}{m^{16}} + \frac{143\lambda^{10} S^1}{m^{17}}\right) g^8 + \mathcal{O}(g^9),
\end{aligned} \tag{3.5.22}$$

$$
\begin{aligned}
\mathcal{F}_{\mathbb{RP}^2} =& \frac{\lambda S}{2m^2}g + \left(\frac{3\lambda S^2}{8m^3} + \frac{\lambda^2 S}{m^4}\right) g^2 + \left(\frac{9\lambda S^2}{4m^5} + \frac{8\lambda^3 S}{3m^6}\right) g^3 \\
&+ \left(\frac{59S^3}{48m^6} + \frac{45\lambda^2 S^2}{4m^7} + \frac{8\lambda^4 S}{m^8}\right) g^4 + \left(\frac{59\lambda S^3}{4m^8} + \frac{105\lambda^3 S^2}{2m^9} + \frac{128\lambda^5 S}{5m^{10}}\right) g^5 \\
&+ \left(\frac{197S^4}{32m^9} + \frac{118\lambda^2 S^3}{m^{10}} + \frac{945\lambda^4 S^2}{4m^{11}} + \frac{256\lambda^6 S}{3m^{12}}\right) g^6 \\
&+ \left(\frac{1773\lambda S^4}{16m^{11}} + \frac{2360\lambda^3 S^3}{3m^{12}} + \frac{2079\lambda^5 S^2}{2m^{13}} + \frac{2048\lambda^7 S}{7m^{14}}\right) g^7 \\
&+ \left(\frac{4775S^5}{128m^{12}} + \frac{19503\lambda^2 S^4}{16m^{13}} + \frac{4720\lambda^4 S^3}{m^{14}} + \frac{9009\lambda^6 S^2}{2m^{15}} + \frac{1024\lambda^8 S}{m^{16}}\right) g^8 \\
&+ \mathcal{O}(g^9),
\end{aligned} \tag{3.5.23}
$$

up to a λ and g independent part. From the tracelessness (3.5.10), we find

$$
\begin{aligned}
\lambda =& \left(-1 + \frac{2}{N}\right) \frac{S}{m} g + \left(-\frac{3}{N} + \frac{12}{N^2}\right) \frac{S^2}{m^4} g^3 + \left(-\frac{1}{N} - \frac{24}{N^2} + \frac{160}{N^3}\right) \frac{S^3}{m^7} g^5 \\
&+ \left(-\frac{3}{4N} - \frac{27}{N^2} - \frac{192}{N^3} + \frac{2688}{N^4}\right) \frac{S^4}{m^{10}} g^7 + \mathcal{O}(g^9) \equiv \widetilde{\lambda}.
\end{aligned} \tag{3.5.24}
$$

Therefore, the effective superpotential is, up to an α independent additive part,

$$
\begin{aligned}
\widetilde{W}_{\text{eff}} =& W_{\text{eff}}|_{\lambda=\widetilde{\lambda}} = \left[N\frac{\partial}{\partial S}\mathcal{F}_{S^2} + 4\mathcal{F}_{\mathbb{RP}^2}\right]\Big|_{\lambda=\widetilde{\lambda}} \\
=& \left(-1 + \frac{4}{N}\right) S^2\alpha + \left(-\frac{1}{3} - \frac{8}{N} + \frac{160}{3N^2}\right) S^3\alpha^2 \\
&+ \left(-\frac{1}{3} - \frac{12}{N} - \frac{256}{3N^2} + \frac{3584}{3N^3}\right) S^4\alpha^3 \\
&+ \left(-\frac{1}{2} - \frac{24}{N} - \frac{352}{N^2} + \frac{33792}{N^4}\right) S^5\alpha^4 + \cdots,
\end{aligned} \tag{3.5.25}
$$

where $\alpha \equiv \frac{g^2}{2m^3}$. This reproduces the result of [23] up to $\mathcal{O}(\alpha^3)$ and extends it further to $\mathcal{O}(\alpha^4)$.

From these examples, the advantage of the present approach over the traceless diagram approach of [23] should be clear. In that approach, one has to evaluate contributing diagrams order by order and evaluating the combinatorics gets very cumbersome. On the other hand, in this traceful approach, there is no issue of

keeping and dropping diagrams, and calculations can be done more systematically. Therefore, being able to reduce the traceless problem to a traceful problem is a great advantage.

3.5.3 Traceless matrix model

We saw that the traceless gauge theory can be solved by the traceful matrix model, not the traceless matrix model. In the following, we argue that the traceless matrix model is not useful in determining the effective superpotential of the traceless gauge theory, $\widetilde{W}_{\text{eff}}$. The relation among traceless and traceful theories, as far as the effective superpotential is concerned, is shown in Fig. 3.1.

$$\begin{array}{ccc} \boxed{\text{traceful gauge theory}} & \leftrightarrow & \boxed{\text{traceful matrix model}} \\ \downarrow & & \downarrow \\ \boxed{\text{traceless gauge theory}} & \not\leftrightarrow & \boxed{\text{traceless matrix model}} \end{array} \tag{3.5.26}$$

Figure 3.1: Relation among traceful and traceless theories

The matrix model loop equation for traceless matter can be derived almost in parallel to the traceless gauge theory loop equation derived in the previous subsection. Again, we replace the projector P with the appropriate traceless version $\widetilde{P}$. For example, in the case of $SU(N)$ adjoint without fundamentals, which was considered in the previous section on the gauge theory side, the loop equation is

$$[W'\tilde{\mathbf{R}}_{S^2}]_- = (\tilde{\mathbf{R}}_{S^2})^2. \tag{3.5.27}$$

Here W is the shifted superpotential defined in (3.5.5), with g_1 defined in (3.5.4) and the gauge theory expectation values replaced by the matrix model expecta-

tion values.

Eq. (3.5.27) is of the same form as the traceful matrix model loop equation, and the first equation of the traceless gauge theory loop equations (3.5.6). Finally, using the equivalence of the traceful gauge theory and matrix model, we conclude that

$$\tilde{\mathbf{R}}(z; g_{p\geq 2}) = \mathbf{R}(z; g_{p\geq 1})\big|_{g_1=\widetilde{g}_1} = R(z; g_{p\geq 1})\big|_{g_1=\widetilde{g}_1} = \widetilde{R}(z; g_{p\geq 2}). \tag{3.5.28}$$

However, what we need to determine $\widetilde{W}_{\text{eff}}$ is $\widetilde{T}$, which we saw in the last subsection to be obtainable from the traceful theory as

$$\begin{aligned}\widetilde{T}(z; g_{p\geq 2}, S, N) &= T(z; g_{p\geq 2}, S, N)\big|_{g_1=\widetilde{g}_1(g_2,g_3,\cdots,S,N)} \\ &= \left[N\frac{\partial}{\partial S}R(z; g_{p\geq 2})\right]\Bigg|_{g_1=\widetilde{g}_1(g_2,g_3,\cdots,S,N)}.\end{aligned} \tag{3.5.29}$$

From the standpoint of the traceless matrix model, the only thing we know is $\tilde{\mathbf{R}} = \widetilde{R} = R|_{g_1=\widetilde{g}_1}$, and we have no information about the g_1 dependence of R. In the framework of the traceless matrix model, there is no way of performing the derivative $\partial/\partial S$ in (3.5.29) before making the replacement $g_1 = \widetilde{g}_1$, because $\widetilde{g}_1$ depends on S also.

Therefore, it is impossible to obtain the effective superpotential for the traceless gauge theory directly, just by using the data from the corresponding traceless matrix model. We really need to invoke the traceful matrix model.

Appendix

3.A Loop equations on the gauge theory side

In this appendix, we are going to calculate the gauge theory loop equations using the approach of [17, 47]. We start with generalized Konishi currents and

corresponding transformations of the fields

$$
\begin{aligned}
J_f &\equiv \mathrm{Tr}\Phi^\dagger e^{V_{\rm adj}} f(\mathcal{W}_\alpha, \Phi) \qquad \therefore \quad \delta\Phi &= f(\mathcal{W}_\alpha, \Phi) \\
J_g &\equiv Q_f^\dagger e^{V_{\rm fund}} g_{ff'}(\Phi) Q_{f'} \qquad \therefore \quad \delta Q_f &= g_{ff'}(\Phi) Q_{f'}
\end{aligned} \tag{3.A.1}
$$

The explicitly written indices on the Q_f's and $g_{ff'}$ are flavor indices, and gauge indices are suppressed. We find the generalized anomaly equations

$$
\begin{aligned}
\bar{D}^2 J_f &= \mathrm{Tr} f(\mathcal{W}_\alpha, \Phi) W'(\Phi) + \tilde{Q} f(\mathcal{W}_\alpha, \Phi) m'(\Phi) Q + \sum_{jklm} A_{jk,lm} \frac{\partial f_{kj}}{\partial \Phi_{lm}} \\
\bar{D}^2 J_g &= 2\tilde{Q} m(\Phi) g(\Phi) Q + \mathrm{Tr} A^{\rm fund} g(\Phi)
\end{aligned} \tag{3.A.2}
$$

and $\bar{D}^2 J_f$ and $\bar{D}^2 J_g$ vanish in the chiral ring.

The field Φ being considered transforms by commutation under gauge transformations, so the elementary anomaly coefficient is the same as the one appearing in [17],

$$
\begin{aligned}
A_{jk,lm} &= \frac{1}{32\pi^2} \left[(\mathcal{W}_\alpha \mathcal{W}^\alpha)_{jm} \delta_{lk} + (\mathcal{W}_\alpha \mathcal{W}^\alpha)_{lk} \delta_{jm} - 2 (\mathcal{W}_\alpha)_{jm} (\mathcal{W}^\alpha)_{lk} \right] \\
&\equiv \frac{1}{32\pi^2} \{ \mathcal{W}_\alpha, [\mathcal{W}^\alpha, e_{ml}] \}_{jk}
\end{aligned} \tag{3.A.3}
$$

where e_{ml} is the basis matrix with the single non-zero entry $(e_{ml})_{jk} = \delta_{mj}\delta_{lk}$. For fields transforming in the fundamental representation we should use

$$
A^{\rm fund}_{jk} = \frac{1}{32\pi^2} (\mathcal{W}_\alpha \mathcal{W}^\alpha)_{jk} \tag{3.A.4}
$$

There is one modification in the treatment of fundamental fields, as compared to the $U(N)$ case studied in [47]. Since the fundamental representation is real for SO and pseudo-real for Sp, the fields Q and $\tilde{Q}$ are not independent; instead, they are related by (3.2.4). This results in the factor of 2 in the second equation in (3.A.2), but otherwise the discussion proceeds as in [47]. In the rest of the Appendix we omit reference to fundamentals.

Next we consider the symmetries of Φ. In equation (3.A.1), $f = \delta\Phi$ must have the same symmetry properties as Φ itself. The tensor field will be taken either symmetric or antisymmetric. We can discuss all four cases in a uniform fashion by using the notation

$$\Phi^T = \begin{cases} \sigma\Phi & \text{for groups } SO(N), \\ \sigma J\Phi J^{-1} & \text{for groups } Sp(N), \end{cases} \tag{3.A.5}$$

and $\sigma = \pm 1$. The gauge field satisfies $\mathcal{W}_\alpha{}^T = -\mathcal{W}_\alpha$ for SO groups, and $\mathcal{W}_\alpha{}^T = -J\mathcal{W}_\alpha J^{-1}$ for Sp groups. As discussed in Subsection 3.2.2, Φ has the property

$$\Phi = P\Phi, \quad \text{or explicitly} \quad \Phi_{ab} = P_{ab,ij}\Phi_{ij} \tag{3.A.6}$$

with the projectors defined in (3.2.14). To ensure that f has the same symmetry as Φ, we should replace $f \to Pf$ in (3.A.2). Specifically, we will take $\delta\Phi$ of the form

$$f^{SO} = P^{SO}\frac{B}{z-\Phi} = \left(\frac{B}{z-\Phi}\right) + \sigma\left(\frac{B}{z-\Phi}\right)^T = \left(\frac{B}{z-\Phi}\right) + \sigma\left(\frac{B^T}{z-\sigma\Phi}\right), \tag{3.A.7}$$

$$f^{Sp} = P^{Sp}\frac{B}{z-\Phi} = \left(\frac{B}{z-\Phi}\right) + \sigma J\left(\frac{B}{z-\Phi}\right)^T J = \left(\frac{B}{z-\Phi}\right) + \sigma\left(\frac{JB^TJ}{z-\sigma\Phi}\right), \tag{3.A.8}$$

with $B = 1$ or $B = \mathcal{W}^2 \equiv \mathcal{W}_\beta\mathcal{W}^\beta$. Using the symmetry of the gauge field and the chiral ring relations, both (3.A.7) and (3.A.8) reduce to

$$f \;=\; \frac{B}{z-\Phi} + \sigma\frac{B}{z-\sigma\Phi}. \tag{3.A.9}$$

Also, to take derivatives with respect to matrix elements[8] correctly we should set

$$\partial_{lm}\Phi_{ab} = P_{lm,ab} \tag{3.A.10}$$

[8]In the case of $U(N)$ of [17], one had $P_{lm,ab} = (e_{lm})_{ab}$ which satisfies $(e_{lm})_{ab}B_{ab} = B_{lm}$, for any matrix B.

Then the tensor field anomaly term becomes

$$A_{jk,lm}\,\partial_{lm} f_{kj} = \frac{1}{32\pi^2}\left[(\mathcal{W}^2)_{jm}\delta_{lk} + \delta_{jm}(\mathcal{W}^2)_{lk} - 2(\mathcal{W}_\alpha)_{jm}(\mathcal{W}^\alpha)_{lk}\right] \times \left[\left(\frac{B}{z-\Phi}\right)_{ra}\left(\frac{1}{z-\Phi}\right)_{bs} P_{kj,rs}P_{lm,ab}\right]. \qquad (3.A.11)$$

After using the projectors (3.2.14), the identity $\mathrm{Tr}\mathcal{W}_\alpha\Phi^k = 0$, the symmetry properties of Φ and $\mathcal{W}_\alpha$, and the chiral ring relations, we find

$$A_{jk,lm}\,\partial_{lm} f_{kj} = \frac{1}{32\pi^2}\left[\left(\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi}\right)\left(\mathrm{Tr}\frac{B}{z-\Phi}\right) + \left(\mathrm{Tr}\frac{1}{z-\Phi}\right)\left(\mathrm{Tr}\frac{\mathcal{W}^2 B}{z-\Phi}\right) + 4k\sigma\,\mathrm{Tr}\frac{\mathcal{W}^2 B}{(z-\Phi)(z-\sigma\Phi)}\right] \qquad (3.A.12)$$

The only difference in (3.A.12) between the two types of gauge groups is that the sign in front of the single trace term is $k=+1$ for SO, and $k=-1$ for Sp. Taking $B=1$ and $B=\mathcal{W}^2$ in (3.A.12) we find

$$0 = \mathrm{Tr}\frac{W'(\Phi)}{z-\Phi} + \sigma\,\mathrm{Tr}\frac{W'(\Phi)}{z-\sigma\Phi} + \frac{2}{32\pi^2}\left[\left(\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi}\right)\left(\mathrm{Tr}\frac{1}{z-\Phi}\right) + 2k\sigma\left(\mathrm{Tr}\frac{\mathcal{W}^2}{(z-\Phi)(z-\sigma\Phi)}\right)\right] \qquad (3.A.13)$$

$$0 = \mathrm{Tr}\frac{\mathcal{W}^2 W'(\Phi)}{z-\Phi} + \sigma\,\mathrm{Tr}\frac{\mathcal{W}^2 W'(\Phi)}{z+\Phi} + \frac{1}{32\pi^2}\left[\left(\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi}\right)\left(\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi}\right)\right] \qquad (3.A.14)$$

Now recall that $W(\Phi)^T = W(\Phi)$ for $SO(N)$, and $W(\Phi)^T = JW(\Phi)J^{-1}$ for $Sp(N)$ since it only appears inside a trace; so

$$\mathrm{Tr}\frac{W'(\Phi)}{z-\sigma\Phi} = \sigma\mathrm{Tr}\frac{W'(\Phi)}{z-\Phi}, \quad \mathrm{Tr}\frac{\mathcal{W}^2 W'(\Phi)}{z-\sigma\Phi} = \sigma\mathrm{Tr}\frac{\mathcal{W}^2 W'(\Phi)}{z-\Phi}. \qquad (3.A.15)$$

The single trace terms have to be treated separately: when $\sigma=-1$,

$$\mathrm{Tr}\frac{\mathcal{W}^2}{z^2-\Phi^2} = \frac{1}{2z}\,\mathrm{Tr}\left[\mathcal{W}^2\left(\frac{1}{z-\Phi}+\frac{1}{z+\Phi}\right)\right] = \frac{1}{z}\,\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi} \qquad (3.A.16)$$

while for $\sigma=+1$, we should use

$$\mathrm{Tr}\frac{\mathcal{W}^2}{(z-\Phi)^2} = -\frac{d}{dz}\,\mathrm{Tr}\frac{\mathcal{W}^2}{z-\Phi}. \qquad (3.A.17)$$

Putting everything together, we find the loop equations written in equation (3.2.17).

3.B Loop equations on the matrix model side

Here we derive the matrix model loop equations for SO/Sp following Seiberg [47], who discussed the $U(N)$ case. Start with the matrix model partition function

$$Z = \int d\mathbf{\Phi} d\mathbf{Q} \exp\left\{-\frac{1}{\mathbf{g}}\left[\mathrm{Tr}[W(\mathbf{\Phi})] + \tilde{\mathbf{Q}}_{\tilde{f}} m_{\tilde{f}f}(\mathbf{\Phi})\mathbf{Q}_f\right]\right\}. \tag{3.B.1}$$

Because the fundamental matter is real for $SO(N)$ and pseudo-real for $Sp(N)$, there is no integration over $\tilde{\mathbf{Q}}$. It is not an independent variable, but related to $\mathbf{Q}$ by Eq. (3.2.4). We will write the symmetry properties of the tensor field $\mathbf{\Phi}$ as

$$\mathbf{\Phi}^T = \begin{cases} \sigma\mathbf{\Phi} & SO(N), \\ \sigma J\mathbf{\Phi}J^{-1} & Sp(N). \end{cases} \tag{3.B.2}$$

where $\sigma = \pm 1$. The matrix $m(\mathbf{\Phi})$ has symmetry properties as given in Eq. (3.2.7).

Now we perform two independent transformations

$$\delta\mathbf{\Phi} = BP\frac{1}{z-\mathbf{\Phi}}, \quad \delta\mathbf{Q}_f = \lambda_{ff'}\frac{1}{z-\mathbf{\Phi}}\mathbf{Q}_{f'} \tag{3.B.3}$$

where B (number) and λ (matrix) are independent and infinitesimal. To make sure that $\delta\mathbf{\Phi}$ has the same symmetry properties as $\mathbf{\Phi}$ itself, we have introduced the appropriate projector P in (3.B.3), see Eq. (3.2.14). The measure in (3.B.1) changes as

$$\begin{aligned} d\mathbf{\Phi} &\rightarrow d\mathbf{\Phi}\, J_{\mathbf{\Phi}} = d\mathbf{\Phi}\,(1+\Delta_{\mathbf{\Phi}}), \\ d\mathbf{Q} &\rightarrow d\mathbf{Q}\, J_{\mathbf{Q}} = d\mathbf{Q}\,(1+\Delta_{\mathbf{Q}}) \end{aligned} \tag{3.B.4}$$

to first order in B and λ, where the corresponding changes in the Jacobians are

$$\begin{aligned} \Delta_{\mathbf{\Phi}} &= BP_{ij,ab}\left(\frac{1}{z-\mathbf{\Phi}}\right)_{ia}\left(\frac{1}{z-\mathbf{\Phi}}\right)_{bj} \\ \Delta_{\mathbf{Q}} &= \left(\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right)(\mathrm{tr}\lambda)\,, \end{aligned} \tag{3.B.5}$$

where tr is a trace over the flavor indices. The classical pieces change by

$$\delta\,\mathrm{Tr}[W(\mathbf{\Phi})] = B\,\mathrm{Tr}\left[W'(\mathbf{\Phi})P\frac{1}{z-\mathbf{\Phi}}\right] = B\,\mathrm{Tr}\frac{W'(\mathbf{\Phi})}{z-\mathbf{\Phi}}. \tag{3.B.6}$$

One can show the second equality using symmetry properties of W: since it only enters Z in the form of the trace, we should take $W(\mathbf{\Phi}^T) = W(\mathbf{\Phi})$ for SO and $W(\mathbf{\Phi}^T) = JW(\mathbf{\Phi})J^{-1}$ for Sp. Similarly,

$$\begin{aligned}\delta(\tilde{\mathbf{Q}}m\mathbf{Q}) &= \tilde{\mathbf{Q}}\left(\frac{\lambda^T m}{z-\sigma\mathbf{\Phi}} + \frac{m\lambda}{z-\mathbf{\Phi}}\right)\mathbf{Q} + B\tilde{\mathbf{Q}}m'\left(P\frac{1}{z-\mathbf{\Phi}}\right)\mathbf{Q} \\ &= 2\tilde{\mathbf{Q}}\frac{m\lambda}{z-\mathbf{\Phi}}\mathbf{Q} + B\tilde{\mathbf{Q}}\frac{m'}{z-\mathbf{\Phi}}\mathbf{Q},\end{aligned} \tag{3.B.7}$$

where we used a similar symmetry property of the matrix m. Finally, with the explicit form of the projectors (3.2.14) we find that in all four cases the statement $\delta Z = 0$ gives two independent loop equations (one for B, and one for λ):

$$\begin{aligned}&\frac{1}{2}\left\langle\left(\mathbf{g}\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right)^2\right\rangle \pm \frac{\sigma}{2}\mathbf{g}\left\langle\mathbf{g}\mathrm{Tr}\frac{1}{(z-\mathbf{\Phi})(z-\sigma\mathbf{\Phi})}\right\rangle \\ &\qquad\qquad = \left\langle\mathbf{g}\mathrm{Tr}\frac{W'(\mathbf{\Phi})}{z-\mathbf{\Phi}}\right\rangle + \mathbf{g}\left\langle\tilde{\mathbf{Q}}\frac{m'(\mathbf{\Phi})}{z-\mathbf{\Phi}}\mathbf{Q}\right\rangle, \\ &\left\langle\mathbf{g}\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right\rangle\delta_{ff'} = 2\left\langle\tilde{\mathbf{Q}}_{\tilde{f}}\frac{m_{\tilde{f}f}(\mathbf{\Phi})}{z-\mathbf{\Phi}}\mathbf{Q}_{f'}\right\rangle,\end{aligned} \tag{3.B.8}$$

for SO and Sp,, respectively. This is Eq. (3.3.8) quoted in Section 3.3.

As it is written, equation (3.B.8) includes all orders in $\mathbf{g}$. The anomaly term in the first equation (3.B.8) factorizes as

$$\left\langle\left(\mathbf{g}\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right)^2\right\rangle = \left\langle\mathbf{g}\mathrm{Tr}\frac{1}{z-\mathbf{\Phi}}\right\rangle^2 \times \left[1+\mathcal{O}(\mathbf{g}^2)\right] \tag{3.B.9}$$

as can be seen from a diagram expansion. With this and the definition of matrix model resolvents (3.3.3) and (3.3.5), we obtain the loop equations (3.3.10).

CHAPTER 4

The string theory prescription

We consider $\mathcal{N} = 1$ supersymmetric $U(N)$, $SO(N)$, and $Sp(N)$ gauge theories, with two-index tensor matter and added tree-level superpotential, for general breaking patterns of the gauge group. By considering the string theory realization and geometric transitions, we clarify when glueball superfields should be included and extremized, or rather set to zero; this issue arises for unbroken group factors of low rank. The string theory results, which are equivalent to those of the matrix model, refer to a particular UV completion of the gauge theory, which could differ from conventional gauge theory results by residual instanton effects. Often, however, these effects exhibit miraculous cancellations, and the string theory or matrix model results end up agreeing with standard gauge theory. In particular, these string theory considerations explain and remove some apparent discrepancies between gauge theories and matrix models in the literature.

4.1 Introduction

Large N topological string duality [56] embedded in superstrings [9, 10] has led to a new perspective on $\mathcal{N} = 1$ supersymmetric gauge theories: that the exact effective superpotential can be efficiently computed by including glueball fields. For example, in a theory with gauge group G, with tree-level superpotential

leading to a breaking pattern

$$G(N) \rightarrow \prod_{i=1}^{K} G_i(N_i) \ , \tag{4.1.1}$$

the dynamics is efficiently encoded in a superpotential $W_{\rm eff}(S_1, \ldots S_K; g_j, \Lambda)$ (g_j are the parameters in $W_{\rm tree}$ and Λ is the dynamical scale). Further, string theory implies [10]

$$W_{\rm eff}(S_i; g_j, \Lambda) = \sum_{i=1}^{K} \left(h_i \frac{\partial \mathcal{F}(S_i)}{\partial S_i} - 2\pi i \tau_i S_i \right) , \tag{4.1.2}$$

with h_i and τ_i the fluxes through A_i and B_i three-cycles in the geometry, as will be reviewed in sect. 4.3. The prepotential $\mathcal{F}(S_i)$ in (4.1.2) is computable in terms of geometric period integrals, which yields [10]

$$\frac{\partial \mathcal{F}(S_i)}{\partial S_i} = S_i \left(\log \left(\frac{\Lambda_i^3}{S_i} \right) + 1 \right) + \frac{\partial}{\partial S_i} \sum_{i_1, \ldots, i_K \geq 0} c_{i_1 \ldots i_K} S_1^{i_1} \cdots S_K^{i_k}, \tag{4.1.3}$$

with coefficients $c_{i_1 \ldots i_K}$ depending on the g_j (but not on the gauge theory scale Λ). In [5] it was shown how planar diagrams of an associated matrix model can also be used to compute (4.1.2) and (4.1.3). Based on the stringy examples, this was generalized in [7] to a more general principle to gain non-perturbative information about the strong coupling dynamics of gauge theories, by extremizing the perturbatively computed glueball superpotential.

There are two aspects to the above statements: first that the glueball fields S_i are the 'right' variables to describe the IR physics, and second that perturbative gauge theory techniques suffice to compute the glueball superpotential. The latter statement has now been proven in two different approaches for low powers of the glueball fields S_i in (4.1.3) [13, 17]. For powers of the glueball fields S_i larger than the dual Coxeter number of the group, an ambiguity sets in for the glueball computation of the coefficients $c_{i_1 \ldots i_K}$ in both of these approaches. The

matrix model provides a natural prescription for how to resolve this ambiguity, essentially by continuing from large N_i. It was argued in [57, 15] that the string geometry / matrix model result (since the string geometry and matrix model results are identical, we refer to them synonymously) has the following meaning: it computes the F-terms for different supersymmetric gauge theories, which can be expressed in terms of $G(N+k|k)$ supergroups. The $W_{\text{eff}}(S_i)$ is independent of k, and the above ambiguity can be eliminated by taking k, and hence the dual Coxeter number, arbitrarily large. The $G(N)$ theory of interest is obtained from the $G(N+k|k)$ theory by Higgsing; but there can be residual instanton contributions to W_{eff} [15], which can lead to apparent discrepancies between the matrix model and gauge theory results. We will somewhat clarify here when such residual instanton effects do, or do not, lead to discrepancies with standard gauge theory results.

There is another, more non-trivial assumption in [7] : the statement that the glueball fields S_i are the 'right' variables in the IR. This assumption was motivated from the string dualities [56] - [10], where geometric transition provide the explanation of why the glueball fields are the natural IR variables: heuristically, $\langle S_i \rangle$ corresponds to confinement. However, this is not quite correct: it also applies to abelian theories, as had been noted in [10] . So the deep explanation of why we should choose certain dynamical S variables remains mysterious.

In this paper, we will uncover the precise prescription for the correct choice of IR variables. This will be done from the string theory perspective, by arguing in which cases there is a geometric transition in string theory. For the general breaking pattern (4.1.1), our prescription for treating the glueball field S_i, corresponding to the factor G_i in (4.1.1), is as follows: *If $h(G_i) > 0$ we include S_i and extremize $W_{\text{eff}}(S_i)$ with respect to it. On the other hand, if $h(G_i) \leq 0$ we*

do not include or extremize S_i, instead we just set $S_i \to 0$. Here we define the generalized dual Coxeter numbers[1]

$$\begin{aligned} h(U(N)) &= N, \\ h(Sp(N)) &= \frac{N}{2} + 1, \\ h(SO(N)) &= N - 2, \end{aligned} \tag{4.1.4}$$

which are generalized in that (4.1.4) applies for all $N \geq 0$. In particular $h(U(1)) = h(Sp(0)) = 1$, so when some G_i factor in (4.1.1) is $U(1)$ or $Sp(0)$, our prescription is to include the corresponding S_i and extremize with respect to it. On the other hand, $h(U(0)) = 0$ and $h(SO(2)) = 0$, so when some G_i factor in (4.1.1) is $U(0)$ or $SO(2)$, our prescription is to just set the corresponding $S_i = 0$ from the outset. (Note that $U(1)$ and $SO(2)$ are treated differently here.)

This investigation was motivated by trying to understand the discrepancies found in [23] for $Sp(N)$ theory with antisymmetric tensor matter, where the superpotentials from the matrix model and gauge theory were found to differ at order h in perturbation theory and beyond. The analysis considered the trivial breaking pattern $Sp(N) \to Sp(N)$ and a single glueball was introduced corresponding to the single unbroken gauge group factor. In [15], various gauge theories including this example were studied, and an explanation for the discrepancies was proposed in terms of the conjecture, mentioned above, that the string theory / matrix model actually computes the superpotential of the large k $G(N+k|k)$ supergroup theories, rather than the ordinary $G(N)$ theory. In this context, the trivial breaking pattern considered in [23] should be understood as $Sp(N) \to Sp(N) \times Sp(0)$, which is completed to $Sp(N+k|k) \to Sp(N+k_1|k_1) \times Sp(k_2|k_2)$. In particular, $Sp(0)$ factors, while

[1] Our convention for symplectic group is such that $Sp(N) \subset SU(N)$ with even N, and hence $Sp(2) \cong SU(2)$.

trivial in standard gauge theory, are non-trivial in the string theory geometry / matrix model context: there can be a residual instanton contribution to the superpotential when one Higgses $Sp(k_2|k_2)$ down to $Sp(0|0) = Sp(0)$, as explicitly seen in [15] for the case of breaking $Sp(0) \to Sp(0)$ with quadratic W_{tree}[2]. Related aspects of "$Sp(0)$" being non-trivial in the string / matrix model context were subsequently discussed in [24, 25, 29].

However, it turns out that one also needs to modify the matrix model side of the computation to take into account the $Sp(0)$ factors. This was found by Cachazo [24], who showed that the loop equations determining $T(z) \equiv \text{Tr}(\frac{1}{z-\Phi})$ and $R(z) \equiv -\frac{1}{32\pi^2}\text{Tr}(\frac{W_\alpha W^\alpha}{z-\Phi})$ for the $Sp(N)$ theory with antisymmetric tensor matter [30, 31] could be related to those of a $U(N+2K)$ gauge theory with adjoint matter, with $Sp(N) \to Sp(N) \times Sp(0)^{K-1}$ mapped to $U(N+2K) \to U(N+2) \times U(2)^{K-1}$. It was thus shown in [24] that vanishing period of $T(z)dz$ through a given cut, corresponding to an $Sp(0)$ factor, does not imply that the cut closes up on shell (aspects of the periods in this theory were also discussed in [26]). This fits with our above prescription that the $Sp(0)$ glueballs should be included and extremized in the string theory / matrix model picture, as would be done for $U(2)$, rather than set to zero, as was originally done in [23]. We stress that we are not yet even discussing whether or not the string theory / matrix model result agrees with standard gauge theory. Irrespective of any comparison with standard gauge theory, the prescription to obtain the actual string theory / matrix model result is as described above (4.1.4). Having obtained that result,

[2] It was suggested in the original version of [15] that such $Sp(0)$ residual instanton contributions could also play a role for the case of cubic and higher order W_{tree} (where they had not yet been fully computed) and could explain the apparent matrix model vs. standard gauge theory discrepancies found in [23]. As we will discuss, we now know that this last speculation was not correct. The corrected proposal of [15] is still that the matrix model computes the superpotential of the $G(N+k|k)$ theory, but where the matrix model side of the computation should be corrected, as we discuss in this paper, to include glueball fields for the $Sp(0)$ factors.

we can now discuss comparisons with standard gauge theory results. As seen in [24], by solving the $U(N+2K)$ loop equations for the present case, this corrected matrix model result now agrees perfectly with standard gauge theory! This will be discussed further here, with all glueball fields S_i included.

This agreement, between the matrix model result and standard gauge theory, is in a sense surprising for this particular theory, in light of the $Sp(k|k)$ description of [15] for the unbroken $Sp(0)$ factors, with the resulting residual instanton contributions to the superpotential. As we will explain later in this paper, the agreement here between matrix models and standard gauge theory is thanks to a remarkable cancellation of the residual instanton effect terms, which could have spoiled the agreement. The cancellation occurs upon summing over the i in (4.1.2) from $i = 1 \ldots K$.

There are similar remarkable cancellations of the "residual instanton contributions" to the superpotential in many other examples, which we will also discuss. In fact, in all cases that we know of, the only cases where the residual instantons do not cancel is when the gauge theory clearly has some ambiguity, requiring a choice of how to define the theory in the UV; string theory / matrix model gives a particular such choice. Examples of such cases is when the LHS of (4.1.1) is itself $U(1)$ or $Sp(0)$ super Yang–Mills, as discussed in [15]. Other examples where the residual instanton contributions do not cancel, is when the superpotential is of high enough order such that not all operators appearing in it are independent, e.g. terms like $\mathrm{Tr}\,\Phi^n$, for a $U(N)$ adjoint Φ, when $n > N$. In standard gauge theory, there are then potential ambiguities involved in reducing such composite operators to the independent operators, since classical operator identities can receive quantum corrections. The residual instanton contributions, which do not cancel generally in these cases, imply specific quantum relations for

these operators, corresponding to the specific UV completion. See [58, 41] for related issues.

The organization of this paper is as follows: In section 4.2 we summarize the gauge theories under consideration. In section 4.3 we review the type IIB string theory construction of these gauge theories. We also discuss maps of the exact superpotentials of Sp and SO theories to those of U theories, generalizing observations of [14, 39, 40, 24, 29]. In section 4.4 we explain, from the string theory perspective in which cases we have a geometric transition. In section 4.5 we consider examples, where the glueball fields S_i of all group factors are correctly accounted for on the matrix model side. The results thereby obtained via matrix models are found to agree with those of standard gauge theory. In many of these examples, this agreement relies on a remarkable interplay of different residual instanton contributions, which sometimes fully cancel. Residual instantons are discussed further in section 4.6, with examples illustrating cases where they do, or do not, cancel. In appendix 4.A, a proof of a general relation between the S^2 and $\mathbb{RP}^2$ contributions to the matrix model free energy is given, and also the matrix model computation of superpotential is presented. In appendix 4.B the gauge theory computation of the superpotential is discussed.

4.2 The gauge theory examples

The specific examples of $\mathcal{N}=1$ supersymmetric gauge theories which we consider, with breaking patterns as in (4.1.1), are as follows:

$$\begin{array}{ll} U(N) \text{ with adjoint } \Phi: & U(N) \to \prod_{i=1}^{K} U(N_i), \\ SO(N) \text{ with adjoint } \Phi: & SO(N) \to SO(N_0) \times \prod_{i=1}^{K} U(N_i), \\ Sp(N) \text{ with adjoint } \Phi: & Sp(N) \to Sp(N_0) \times \prod_{i=1}^{K} U(N_i), \\ SO(N) \text{ with symmetric } S: & SO(N) \to \prod_{i=1}^{K} SO(N_i), \\ Sp(N) \text{ with antisymmetric } A: & Sp(N) \to \prod_{i=1}^{K} Sp(N_i), \\ U(N) \text{ with } \Phi + S + \widetilde{S}: & U(N) \to SO(N_0) \times \prod_{i=1}^{K} U(N_i), \\ U(N) \text{ with } \Phi + A + \widetilde{A}: & U(N) \to Sp(N_0) \times \prod_{i=1}^{K} U(N_i). \end{array} \tag{4.2.1}$$

For $U(N)$ with adjoint Φ, the tree-level superpotential is taken to be

$$W_{\text{tree}} = \text{Tr}[W(\Phi)], \qquad W(x) = \sum_{j=1}^{K+1} \frac{g_j}{j} x^j, \tag{4.2.2}$$

with K potential wells. In the classical vacua, with breaking pattern as in (4.2.1), Φ has N_i eigenvalues equal to the root a_i of

$$W'(x) = \sum_{j=1}^{K+1} g_j x^{j-1} \equiv g_{K+1} \prod_{i=1}^{K} (x - a_i), \tag{4.2.3}$$

with $\sum_{i=1}^{K} N_i = N$. For $SO(N)$ with symmetric tensor or $Sp(N)$ with antisymmetric tensor we take

$$W_{\text{tree}} = \frac{1}{2}\text{Tr}W(S), \qquad \text{or} \qquad W_{\text{tree}} = \frac{1}{2}\text{Tr}W(A), \tag{4.2.4}$$

respectively, where $W(x)$ is as in (4.2.2), the factor of $\frac{1}{2}$ is for convenience, because the eigenvalues of S or A appear in pairs, and the indices are contracted with δ^a_b for $SO(N)$ or J^a_b for $Sp(N)$. For $SO/Sp(N)$[3] with adjoint matter, the tree-level

[3]We often write $SO(N)$ and $Sp(N)$ simply as $SO/Sp(N)$. When we use a "$\pm$" sign for $SO/Sp(N)$, it means "$+$" for $SO(N)$ and "$-$" for $Sp(N)$. Similarly "$\mp$" means "$-$" for $SO(N)$ and "$+$" for $Sp(N)$.

superpotential is

$$W_{\text{tree}} = \frac{1}{2}\text{Tr}[W(\Phi)], \qquad W(x) = \sum_{j=1}^{K+1} \frac{g_{2j}}{2j} x^{2j}, \tag{4.2.5}$$

since all Casimirs of the adjoint Φ are even, and the $\frac{1}{2}$ is again for convenience because the eigenvalues appear in pairs. Φ's eigenvalues sit at the zeros of

$$W'(x) = \sum_{j=1}^{K+1} g_{2j} x^{2j-1} \equiv g_{2K+2} x \prod_{i=1}^{K} (x^2 - a_i^2). \tag{4.2.6}$$

The breaking pattern in (4.2.1) has N_0 eigenvalues of Φ equal to zero, and N_i pairs at $\pm a_i$, so $N = N_0 + 2\sum_{i=1}^{K} N_i$ for $SO/Sp(N) \to SO/Sp(N_0) \times \prod_{i=1}^{K} U(N_i)$ (with the convention $Sp(2) \cong SU(2)$).

The next to last example in (4.2.1) is the $\mathcal{N} = 2$ $U(N)$ theory with a matter hypermultiplet in the two-index symmetric tensor representation, breaking $\mathcal{N} = 2$ to $\mathcal{N} = 1$ by a superpotential as in (4.2.2):

$$W = \sum_{j=1}^{K+1} \frac{g_j}{j} \text{Tr}\Phi^j + \sqrt{2}\,\text{Tr}\widetilde{S}\Phi S. \tag{4.2.7}$$

In addition to the possibility of Φ's eigenvalues sitting in any of the K critical points $W'(x)$ analogous to (4.2.3), there is a vacuum where N_0 eigenvalues sits at $\phi = 0$, with $\langle S\widetilde{S}\rangle \neq 0$, breaking $U(N_0) \to SO(N_0)$. The last example in (4.2.1) is the similar theory where the $\mathcal{N} = 2$ hypermultiplet is instead in the antisymmetric tensor representation A, rather than the symmetric tensor S. These last two classes of examples were considered in [50, 46, 27].

In all of these theories, the low energy superpotential is of the general form

$$W_{\text{low}}(g_j, \Lambda) = W_{\text{cl}}(g_j) + W_{\text{gc}}(\Lambda_i) + W_H(g_j, \Lambda). \tag{4.2.8}$$

$W_{\text{cl}}(g_j)$ is the classical contribution (evaluating W_{tree} in the appropriate minima). $W_{\text{gc}}(\Lambda_i)$ is the gaugino condensation contribution in the unbroken gauge groups

of (4.1.1),

$$W_{\rm gc}(\Lambda_j) = \sum_{i=1}^{K} h_i e^{2\pi i n_i/h_i} \Lambda_i^3, \tag{4.2.9}$$

where $h_i = C_2(G_i)$ is the dual Coxeter number of the group G_i in (4.1.1), with the phase factors associated with the $\mathbb{Z}_{2h_i} \to \mathbb{Z}_2$ chiral symmetry breaking of the low-energy G_i gaugino condensation. The scales Λ_i are related to Λ by threshold matching for the fields which got a mass from $W_{\rm tree}$ and the breaking (4.1.1); some examples, with breaking patterns as in (4.2.1), are as follows. For $U(N)$ with adjoint Φ:

$$\Lambda_i^{3N_i} = \Lambda^{2N} W''(a_i)^{N_i} \prod_{j\neq i} m_{W_{ij}}^{-2N_j} = g_{K+1}^{N_i} \Lambda^{2N} \prod_{j\neq i} (a_j - a_i)^{N_i - 2N_j}. \tag{4.2.10}$$

For $SO/Sp(N)$ with adjoint, breaking $SO/Sp(N) \to SO/Sp(N_0) \times \prod_{i=1}^{K} U(N_i)$,

$$\begin{aligned} \Lambda_0^{3(N_0\mp 2)} &= g_{2K+2}^{N_0\mp 2} \Lambda^{2(N\mp 2)} \prod_{i=1}^{K} a_i^{2(N_0\mp 2)-4N_i} \\ \Lambda_i^{3N_i} &= 2^{-N_i} g_{2K+2}^{N_i} \Lambda^{2(N\mp 2)} a_i^{-2(N_0\mp 2)} \prod_{j\neq i}^{K} (a_i^2 - a_j^2)^{N_i - 2N_j}. \end{aligned} \tag{4.2.11}$$

Note that, although this relation looks very similar for $SO(N)$ and $Sp(N)$, its meaning is different for two groups in gauge theory due to the index of the embedding of $U(N_i)$; for $SO(N)$, the $U(N_i)$ one-instanton factor $\Lambda_i^{b(U(N_i))}$, $b(U(N_i)) = 3N_i$, is related to the $SO(N)$ one-instanton factor $\Lambda^{b(SO(N))}$, $b(SO(N)) = 2(N-2)$. On the other hand, for $Sp(N)$, it is related to the $Sp(N)$ *two*-instanton factor $\Lambda^{2b(Sp(N))}$, $b(Sp(N)) = N+2$. For $SO(N)$ with symmetric tensor,

$$\Lambda_i^{3(N_i-2)} = g_{K+1}^{N_i+2} \Lambda^{2N-8} \prod_{j\neq i=1}^{K} (a_i - a_j)^{N_i+2-2N_j}. \tag{4.2.12}$$

For $Sp(N)$ with antisymmetric A:

$$\Lambda_i^{3(N_i+1)} = g_{K+1}^{N_i-1} \Lambda^{2N+4} \prod_{j\neq i}^{K} (a_i - a_j)^{N_i-1-2N_j}. \tag{4.2.13}$$

Finally, the term $W_H(g_j, \Lambda)$ in (4.2.8) are additional non-perturbative contributions, which can be regarded as coming from the massive, broken parts of the gauge group.

In the description with the glueballs S_i integrated in, as in (4.1.2), the gaugino condensation contribution comes from the first term in (4.1.3):

$$W_{\rm gc}(S_i, \Lambda) = \sum_{i=1}^{K} h_i S_i \left(\log \left(\frac{\Lambda_i^3}{S_i} \right) + 1 \right) \ , \tag{4.2.14}$$

and $W_H(g_i, \Lambda)$ comes from the last terms in (4.1.3), upon integrating out the S_i. When the minima a_i of the superpotential are widely separated, the contributions W_H from these last terms are subleading as compared with $W_{\rm gc}$. As in (4.1.2), the full glueball superpotential (4.1.2) can be computed via the string theory geometric transition, in terms of certain period integrals [9, 10], as will be reviewed in the next section, or via the matrix models. In that context, the term $W_{\rm gc}(S_i, \Lambda_i)$ comes from the integration measure (as is also natural in field theory, since it incorporates the $U(1)_R$ anomaly) and $W_H(S_i, g_j)$ can be computed perturbatively [5, 7]. The perturbative computation of $W_H(S_i, g_j)$ can also be understood directly in the gauge theory [13], up to ambiguities in terms S^n with $n > h = C_2(G)$. The string theory/ matrix model constructions yield a specific way of resolving these ambiguities, which correspond to a particular UV completion of the gauge theory [57, 15].

As discussed in the introduction, our present interest will be in analyzing this circle of ideas when some of the gauge group factors in (4.1.1) are of low rank, or would naively appear to be trivial, e.g. $U(1)$, $U(0)$, $SO(2)$, $SO(0)$, and $Sp(0)$.

4.3 Geometric transition of $U(N)$ and $SO/Sp(N)$ theories

In this section we briefly review the type IIB geometric engineering of relevant $U(N)$ and $SO/Sp(N)$ theories, and their geometric transition.

4.3.1 $U(N)$ with adjoint and $W_{\mathbf{tree}} = \mathrm{Tr} \sum_{j=1}^{K+1} \frac{g_i}{j} \Phi^j$

The Calabi–Yau 3-fold relevant to this theory [10, 59] is the non-compact A_1 fibration

$$W'(x)^2 + y^2 + u^2 + v^2 = 0 \ . \tag{4.3.1}$$

This fibration has K conifold singularities at the critical points of $W(x)$, i.e. at $W'(x) = 0$. Near each of the singularities, the geometry (4.3.1) is the same as the usual conifold $x'^2 + y^2 + u^2 + v^2 = 0$, which is topologically a cone with base $S^2 \times S^3$.

The singularities can be resolved by blowing up a 2-sphere $S^2 = \mathbb{P}^1$ at each singularity. We can realize the $U(N)$ gauge theory with adjoint matter and superpotential (4.2.2) in type IIB superstring theory compactified on this resolved geometry, with N D5-branes partially wrapping the K $\mathbb{P}^1$'s; two dimensions of the D5-brane worldvolume wrap the $\mathbb{P}^1$'s and the remaining four dimensions fill the flat Minkowski space. The gauge theory degrees of freedom correspond to the open strings living on these D5-branes. The classical supersymmetric vacuum is obtained by distributing the N_i D5-branes over the i-th 2-sphere $\mathbb{P}^1_i$ with $i = 1, \cdots, K$. The corresponding breaking pattern of the gauge group is as in (4.2.1): $U(N) \to \prod_{i=1}^{K} U(N_i)$.

At low energy, the gauge theory confines (when $N_i > 1$), each $U(N_i)$ factor developing nonzero vev of the glueball superfield S_i. In string theory this is

described by the geometric transition [56, 9, 10] in which the resolved conifold geometry with $\mathbb{P}^1$'s wrapped by D5-branes is replaced by a deformed conifold geometry

$$W'(x)^2 + f_{K-1}(x) + y^2 + u^2 + v^2 = 0 \ , \tag{4.3.2}$$

where $f_{K-1}(x)$ is a polynomial of degree $K-1$ in x and parametrizes the deformation. After the geometric transition, each 2-sphere $\mathbb{P}^1_i$ wrapped by N_i D5-branes is replaced by a 3-sphere A_i with 3-form RR flux through it:

$$\oint_{A_i} H = N_i \ , \tag{4.3.3}$$

where $H = H_{RR} + \tau H_{NS}$ and $\tau = C^{(0)} + ie^{-\Phi}$ is the complexified coupling constant of type IIB theory. We define the periods of the Calabi–Yau geometry (4.3.2) by

$$S_i \equiv \frac{1}{2\pi i}\oint_{A_i} \Omega, \qquad \Pi_i \equiv \int_{B_i}^{\Lambda_b} \Omega, \tag{4.3.4}$$

where Ω is the holomorphic 3-form, B_i is the noncompact 3-cycle dual to the 3-cycle A_i, and Λ_b is a cutoff needed to regulate the divergent B_i integrals. The IR cutoff Λ_b is to be identified with the UV cutoff of the 4d gauge theory. The set of variables S_i measure the size of the blown up 3-spheres, and can be used to parametrize the deformation in place of the K coefficients of the polynomial $f_{K-1}(x)$.

The dual theory after the geometric transition is described by a 4d, $\mathcal{N} = 1$ $U(1)^K$ gauge theory, with K $U(1)$ vector superfields V_i and K chiral superfields S_i. If not for the fluxes, this theory would be $\mathcal{N} = 2$ supersymmetric $U(1)^K$, with (V_i, S_i) the $\mathcal{N} = 2$ vector super-multiplets, S_i being the Coulomb branch moduli. This low-energy $U(1)^K$ theory has a non-trivial prepotential $\mathcal{F}(S_i)$ and the dual periods Π_i in (4.3.4) can be written as

$$\Pi_i(S) = \frac{\partial \mathcal{F}}{\partial S_i}. \tag{4.3.5}$$

Without the fluxes, this prepotential can be understood as coming from integrating out D3 branes which wrap the A_i cycles, and are charged under the low-energy $U(1)$'s [60].

The effect of the added fluxes is to break $\mathcal{N}=2$ supersymmetry to $\mathcal{N}=1$ by the added superpotential [61, 62]

$$W_{\text{flux}} = \int H \wedge \Omega = \sum_i \left(\oint_{A_i} H \int_{B_i}^{\Lambda_b} \Omega - \int_{B_i}^{\Lambda_b} H \oint_{A_i} \Omega \right) . \tag{4.3.6}$$

In the present $U(N)$ case, (4.3.3) and (4.3.4) gives

$$W_{\text{flux}} = \sum_{i=1}^{K} (N_i \Pi_i - 2\pi i \alpha S_i) \ , \tag{4.3.7}$$

where

$$\int_{B_i}^{\Lambda_b} H = \alpha \tag{4.3.8}$$

is the 3-form NS flux through the 3-cycle B_i and identified with the bare coupling constant of the gauge theory by

$$2\pi i \alpha = \frac{8\pi^2}{g_b^2} = V. \tag{4.3.9}$$

where V is the complexified volume of the $\mathbb{P}^1$'s. The $\mathcal{N}=1$ $U(1)^K$ vector multiplets V_i remain massless, but the S_i now have a superpotential, which fixes them to sit at discrete vacuum expectation values, where they are massive. The fields S_i are identified with the glueballs on the gauge theory side.

The superpotential $W_{\text{flux}}(S_i)$ is the full, exact, effective superpotential in (4.1.2). As can be verified by explicit calculations [9, 10], the leading contribution to (4.3.7) is always of the form

$$W_{\text{flux}} \sim \sum_{i=1}^{K} [N_i S_i (1 - \ln(S_i/\Lambda_i^3)) - 2\pi i \alpha S_i], \tag{4.3.10}$$

where Λ_i is related to the scale Λ_b via precisely the relation (4.2.10). This leading term (4.3.10) is the gaugino condensation part of the superpotential, as in (4.1.3).

4.3.2 SO and Sp theories

The string theory construction of $SO/Sp(N)$ theories can be obtained from the above $U(N)$ construction, by orientifolding the geometry before and after the geometric transition by a certain $\mathbb{Z}_2$ action. The geometric construction of $SO/Sp(N)$ theory with adjoint was discussed in [63, 64, 39, 65], and in that case the invariance of the geometry (4.3.2) under the $\mathbb{Z}_2$ action requires that the polynomial $W(x)$ be even. The geometric construction of $SO/Sp(N)$ theory with symmetric/antisymmetric tensor was studied in [66, 67, 29].

In the classical vacuum of the "parent" $U(2N)$ theory, the gauge group is broken into a product of $U(N_i)$ groups. When a $U(N_i)$ factor is identified with another $U(N_i)$ by the $\mathbb{Z}_2$ orientifold action, they lead to a single $U(N_i)$ factor. When a $U(N_i)$ factor is mapped to itself by the $\mathbb{Z}_2$ orientifold action, it becomes an $SO(N_i)$ or $Sp(N_i)$, depending on the charge of the orientifold hyperplane. As a result, the classical vacuum of the "daughter" $SO/Sp(N)$ theory has gauge group broken as in (4.2.1), depending on whether the theory is SO or Sp with adjoint, or SO with symmetric tensor, or Sp with antisymmetric tensor:

$$SO/Sp(N) \to \prod_i G_i(N_i), \qquad G_i = U, SO, \text{ or } Sp. \tag{4.3.11}$$

The 3-form RR fluxes from orientifold hyperplanes makes an additional contribution to the superpotential (4.3.6), and the flux superpotential can be written as

$$W_{\text{flux}} = \sum_i [\widehat{N}_i \Pi_i(S_i) - 2\pi i \eta_i \alpha S_i], \tag{4.3.12}$$

where

$$\widehat{N}_i = \begin{cases} N_i & G_i = U(N_i), \\ N_i/2 \mp 1 & G_i = SO/Sp(N_i), \end{cases}$$
$$\eta_i = \begin{cases} 1 & G_i = U, \\ 1/2 & G_i = SO/Sp. \end{cases} \tag{4.3.13}$$

$\widehat{N}_i$ is the net 3-form RR flux through the A_i cycle. For $U(N_i)$ and $Sp(N_i)$, $\widehat{N}_i$ in (4.3.13) is the dual Coxeter number (4.1.4), while for $SO(N_i)$ it is half[4] the dual Coxeter number (4.1.4). The 1/2 in (4.3.13) is because the integration over the A_i cycles should be halved due to the $\mathbb{Z}_2$ identification.

4.3.3 Relations between SO/Sp theories and $U(N)$ theories

The result (4.3.12), with (4.3.13), gives the exact superpotential of the SO/Sp theories in terms of the same periods S_i and $\Pi(S_i)$ as an auxiliary U theory. This was first noted in [24] at the level of the Konishi anomaly equation as a map between the resolvents of Sp theory with antisymmetric matter and U theory with adjoint matter. In [29], it was generalized to the map between the resolvents of SO/Sp theories with two-index tensor matter and U theory with adjoint matter, and string theory interpretation was discussed. In this subsection we will derive this map from the string theory perspective using the flux superpotential (4.3.12). Furthermore, we will clarify the relation of the superpotential and the scale of the SO/Sp theories to those of the U theory. The map between resolvents can be derived from these results. For $Sp(N)$ theory with an antisymmetric tensor,

[4] So we get h replaced with $h/2$ for SO groups in (4.2.9). While one could absorb the overall factor of 2 into the definition of Λ, the number of vacua should be h whereas here we apparently get $h/2$ for SO groups. This is because the we don't see spinors or the $\mathbb{Z}_2$ part of the center which acts on them; it's analogous to $U(2N)$ being restricted to vacua with confinement index 2.

the scale relation was obtained in a different way in [24].

As a first example, consider $SO(N)$ with an adjoint, with the breaking pattern as in (4.2.1). The geometric transition result (4.3.12) and (4.3.13) implies that the exact superpotential is the same as for the $U(N-2)$ theory with adjoint, with breaking pattern map

$$SO(N) \to SO(N_0) \times \prod_{i=1}^{K} U(N_i) \quad \Longleftrightarrow \quad U(N-2) \to U(N_0-2) \times \prod_{i=1}^{K} U(N_i)^2. \tag{4.3.14}$$

The map between the superpotential is

$$W_{\text{exact}}^{SO(N)} = \frac{1}{2} W_{\text{exact}}^{U(N-2)}. \tag{4.3.15}$$

The $SO(N)$ scale matching relation (4.2.11) is compatible with the map (4.3.14), since (4.2.10) for the theory (4.3.14) reproduces (4.2.11).

Likewise, $Sp(N)$ with adjoint has the same exact superpotential as for the $U(N+2)$ theory with adjoint, with

$$\begin{gathered} Sp(N) \to Sp(N_0) \times \prod_{i=1}^{K} U(N_i) \quad \Longleftrightarrow \quad U(N+2) \to U(N_0+2) \times \prod_{i=1}^{K} U(N_i)^2, \\ W_{\text{exact}}^{Sp(N)} = \frac{1}{2} W_{\text{exact}}^{U(N+2)}. \end{gathered} \tag{4.3.16}$$

The $Sp(N)$ scale matching relations (4.2.11) follow from the $U(N)$ matching relations (4.2.10) with the replacement (4.3.16), with the understanding that the $U(N+2)$ and $U(N_0+2)$ one-instanton factors correspond to the $Sp(N)$ and $Sp(N_0)$ two-instanton factors; this is related to the index of the embedding mentioned after (4.2.11), and is accounted for by dividing the $U(N+2)$ superpotential by two, as above.

Next consider $Sp(N)$ with antisymmetric tensor A and breaking pattern as in (4.2.1). The geometric transition result (4.3.12) and (4.3.13) implies that the

exact superpotential is the same as for the $U(N+2K)$ theory with adjoint and breaking pattern

$$Sp(N) \to \prod_{i=1}^{K} Sp(N_i) \quad \Longleftrightarrow \quad U(N+2K) \to \prod_{i=1}^{K} U(N_i+2). \tag{4.3.17}$$

In the present case, comparing the matching relations (4.2.13) for $Sp(N)$ with symmetric tensor with the matching relations (4.2.10) for $U(N+2K)$ with adjoint, for the mapping as in (4.3.17) requires that the scales of the original $Sp(N)$ on the LHS of (4.3.17) and the $U(N+2K)$ on the RHS of (4.3.17) be related as

$$\Lambda_{U(N+2K)}^{N+2K} = g_{K+1}^{-2} \Lambda_{Sp(N)}^{N+4}. \tag{4.3.18}$$

Then the Λ_i of the unbroken groups on both sides of (4.3.17) coincide, with the understanding that the $U(N_i+2)$ one-instanton factors correspond to the $Sp(N_i)$ two-instanton factors, as above. The map between the superpotential is

$$\begin{aligned} W_{\text{exact}}^{Sp(N)} &= \frac{1}{2}[W_{\text{exact}}^{U(N+2K)} - \Delta W_{\text{cl}}], \qquad \Delta W_{\text{cl}} = 2\sum_{i=1}^{K} W(a_i), \\ \text{i.e. writing} \quad W_{\text{exact}} &= W_{\text{cl}} + W_{\text{quant}}, \qquad W_{\text{quant}}^{Sp(N)} = \frac{1}{2} W_{\text{quant}}^{U(N+2K)}. \end{aligned} \tag{4.3.19}$$

Note that, in order for the superpotentials on the two sides of (4.3.17) to fully coincide, one must compensate the classical mismatch ΔW_{cl}, since each well is occupied by two additional eigenvalues in the theory on the RHS of (4.3.17). In the string theory geometric transition realization, this constant shift, which is independent of N, Λ, and the glueball fields S_i, is most naturally interpreted as an additive shift of the superpotential on the Sp side, which can be regarded as coming from the orientifold planes both before and after the transitions. The classical shift of ΔW_{cl} leads to slightly different operator expectation values (as computed via $W_{\text{eff}}(g_p, \Lambda)$ as the generating function) between the Sp and U theory, as was seen in the example of [24]. Also, writing the map as in (4.3.17), we

want the vacuum with confinement index 2 [24]. We could equivalently replace the RHS of (4.3.17) with $U(N/2+K) \to \prod_{i=1}^{K} U(N_i/2+1)$, in which case we would not have to divide by 2 in (4.3.19).

Likewise, $SO(N)$ with symmetric tensor S has exact superpotential related to that of a $U(N-2K)$ theory with adjoint as

$$
\begin{gathered}
SO(N) \to \prod_{i=1}^{K} SO(N_i) \quad \Longleftrightarrow \quad U(N-2K) \to \prod_{i=1}^{K} U(N_i-2), \\
W_{\text{exact}}^{SO(N)} = \frac{1}{2}[W_{\text{exact}}^{U(N-2K)} + \Delta W_{\text{cl}}], \qquad \Delta W_{\text{cl}} = 2\sum_{i=1}^{K} W(a_i).
\end{gathered}
\tag{4.3.20}
$$

Comparing the matching relations (4.2.12) for $SO(N)$ with symmetric tensor with those of (4.2.10) for $U(N-2K)$ with adjoint requires that the scales of the original $SO(N)$ on the LHS of (4.3.20) and those of the $U(N-2K)$ theory on the RHS be related as

$$
\Lambda_{U(N-2K)}^{2(N-2K)} = g_{K+1}^{4} \Lambda_{SO(N)}^{2N-8}. \tag{4.3.21}
$$

Then the Λ_i of the unbroken groups on both sides of (4.3.20) coincide. Again, in order for the superpotentials on the two sides of (4.3.20) to fully coincide, one must correct for the classical mismatch ΔW_{cl} coming from the fact that the $U(N-2K)$ theory has two fewer eigenvalues in each well.

In appendix 4.A we will discuss these relations from the matrix model viewpoint. In this context, the relation relevant for (4.3.14) and (4.3.16) was conjectured in [14, 39] based on explicit diagrammatic calculations, and it was proven for the case of unbroken gauge group $N_0 = N$ in [40]. This will be generalized in Appendix 4.A to all breaking patterns. Likewise, the matrix model relation relevant for (4.3.17) and (4.3.20) will be proven in the appendix; this is a generalization of the connection found in [24] for the theories in (4.3.17) and (4.3.20).

4.4 String theory prescription for low rank

The discussion of the previous section applies for all $N_i \geq 0$. We now discuss under which circumstances one expects a transition in string theory, where S_i^3's grows, and therefore an effective glueball field S_i should be included in the superpotential. Whether or not there is a geometric transition in string theory is a local question, so each S_i^3 can be studied independently. Near any S_i^3 the local physics is just a conifold singularity, so we only need to consider the case of a conifold singularity.

4.4.1 Physics near a conifold singularity

As we saw, $U(N)$, $SO/Sp(N)$ gauge theory can be realized in type IIB theory as the open string theory living on the D5-branes partially wrapped on the exceptional $\mathbb{P}^1$ of a resolved conifold geometry. There is a $\mathbb{P}^1$ associated with each critical point of the polynomial $W(x)$. By the geometric transition duality [9, 10], this gauge theory is dual to the closed string theory in the deformed conifold geometry where the $\mathbb{P}^1$'s have been blown down and S^3's are blown up instead.

Let us focus on one $\mathbb{P}^1$ with $N \geq 0$ D5-branes wrapping it. This corresponds to focusing on one critical point on the gauge theory side. We allow $N = 0$, which corresponds to an unoccupied critical point. In the neighborhood, the geometry after the geometric transition is approximately a deformed conifold $x^2 + y^2 + z^2 + w^2 = \mu$ with a blown up S^3. The low energy degrees of freedom in the four-dimensional theory are the $\mathcal{N} = 1$ $U(1)$ photon vector superfield V and the $\mathcal{N} = 1$ chiral superfield S. The chiral superfield S is neutral under the $U(1)$, and can be thought of as in the adjoint representation of the $U(1)$. The bosonic

component of S is proportional to μ and measures the size of the S^3.

First, consider the case *without* fluxes. Then the closed string theory has $\mathcal{N} = 2$ and there is one $\mathcal{N} = 2$ $U(1)$ vector multiplet (V, S). It is known that as the size of the S^3 goes to zero there appears an extra massless degree of freedom [60], which corresponds to the D3-brane wrapping the S^3. The mass of the wrapped BPS D3-brane is proportional to the area, S, of the S^3, so the mass becomes zero as the S^3 shrinks to zero, i.e. as $S \to 0$. This extra degree of freedom is described as an $\mathcal{N} = 2$ hypermultiplet charged under the $U(1)$ (of V). Let us write this hypermultiplet in $\mathcal{N} = 1$ language as $(Q, \widetilde{Q})$, where Q and $\widetilde{Q}$ are both $\mathcal{N} = 1$ chiral superfields with opposite $U(1)$ charges. The $\mathcal{N} = 2$ supersymmetry requires the superpotential

$$W_Q = \sqrt{2}\, Q\widetilde{Q}S, \tag{4.4.1}$$

which indeed incorporates the above situation that the $Q, \widetilde{Q}$ become massless as $S \to 0$. The D-flatness is

$$|Q|^2 - |\widetilde{Q}|^2 = 0, \tag{4.4.2}$$

and the F-flatness is

$$QS = \widetilde{Q}S = Q\widetilde{Q} = 0. \tag{4.4.3}$$

The only solution to these is

$$Q = \widetilde{Q} = 0, \qquad S : \text{any}, \tag{4.4.4}$$

which just means that $S \sim \mu$ is a modulus.

Now let us come back to the case with the fluxes. As reviewed in the last section, the fluxes give rise to a superpotential (4.3.6) which breaks $\mathcal{N} = 2$ to $\mathcal{N} = 1$ [61, 62]. As in (4.3.12), the local flux superpotential contribution is

$$W_{\text{flux}}(S) \simeq \widehat{N}S[1 - \ln(S/\Lambda^3)] - 2\pi i\eta\alpha S, \tag{4.4.5}$$

where we just keep the leading order term in (4.3.12), as in (4.3.10), with

$$\widehat{N} = \begin{cases} N & U(N), \\ N/2 \mp 1 & SO/Sp(N), \end{cases} \qquad \eta = \begin{cases} 1 & U(N), \\ 1/2 & SO/Sp(N). \end{cases} \tag{4.4.6}$$

The scale Λ is written in terms of the bare coupling Λ_b and the coupling constants in the problem, as before, and $2\pi i\alpha$ is related to the bare gauge coupling by (4.3.9).

In the following, we discuss the cases with $\widehat{N} = 0$, $\widehat{N} > 0$ and $\widehat{N} < 0$ in order.

- **$\hat{N} = 0$ case**

  In this case, the total superpotential is simply the sum of (4.4.1) and (4.4.5):

$$W = \sqrt{2}\, Q\widetilde{Q}S - 2\pi i\eta\alpha S. \tag{4.4.7}$$

  The only solution to the equation of motion is

$$|Q|^2 = |\widetilde{Q}|^2, \qquad Q\widetilde{Q} = \frac{2\pi i\eta\alpha}{\sqrt{2}}, \qquad S = 0. \tag{4.4.8}$$

  This is consistent with the fact that α is proportional to the volume of the $\mathbb{P}^1$, and the D3-brane condensation $\langle Q\widetilde{Q}\rangle$ corresponds to the size of the $\mathbb{P}^1$. Furthermore, since $\langle S\rangle = 0$, the superpotential vanishes: $W = 0$. Therefore, for $\widehat{N} = 0$, i.e. for $U(0)$ and $SO(2)$, geometric transition does not take place and we should set the corresponding glueball field $S \to 0$ from the beginning.

- **$\hat{N} > 0$ case**

  In this case, there is a net RR flux through the A-cycle: $\oint_A H = \widehat{N}$. This means that the D3-brane hypermultiplet $(Q, \widetilde{Q})$ is infinitely massive, because the RR flux will induce $\widehat{N}$ units of fundamental charge on the D3-brane. Since the D3-brane is wrapping a compact space S^3, the fundamental

charge on it should be canceled by $\widehat{N}$ fundamental strings attached to it. Those fundamental strings extend to infinity and thus cost infinite energy[5]. Therefore, we can forget about $Q, \widetilde{Q}$ in this case, and the full superpotential is given just by the flux contribution (4.4.5). The equation of motion gives

$$S^{\widehat{N}} \simeq \Lambda^{3\widehat{N}} e^{-2\pi i \eta \alpha}, \tag{4.4.9}$$

which corresponds to the confining vacua of the gauge theory. Note that this case includes $U(1)$ and $Sp(0)$; these theories have a dual confining description. This may sound a little paradoxical, but is related to the fact that the string theory computes not for the standard $G(N)$ gauge theory but the associated $G(N+k|k)$ higher rank gauge theory, which is confining and differs from standard $U(1)$ and $Sp(0)$ due to residual instanton effects [15].

- **$\hat{N} < 0$ case**

  In this case, the same argument as the $\widehat{N} > 0$ case tells us that we should not include the D3-brane fields $Q, \widetilde{Q}$. Hence the superpotential is just the flux part (4.4.5), which again leads to

$$S^{\widehat{N}} \simeq \Lambda^{3\widehat{N}} e^{-2\pi i \eta \alpha}. \tag{4.4.10}$$

However, now (4.4.10) is physically unacceptable, since S diverges in the weak coupling limit where the bare volume of $\mathbb{P}^1$ becomes $V = 2\pi i \alpha \to \infty$ ($g_b \to 0$) — i.e. taking $\mathbb{P}^1$ large would lead to S^3 also being large, which does not make sense geometrically. The resolution is that S cannot be a good variable: the S^3 does not actually blow up, and S should be set to zero, $S \to 0$, also for this case. Though S is set to zero, the non-zero flux can lead to a non-zero superpotential contribution $W_{\text{flux}} = \widehat{N}\Pi(S \to 0)$.

[5]This phenomenon is the same as that observed in the context of AdS/CFT [68, 69, 70].

Note that the above result concerning the sign of $\widehat{N}$ does not mean the gauge theory prefers D5-branes to anti-D5-branes; it just means that one should choose the sign of the NS flux (i.e. the sign of $2\pi i\alpha$) appropriately. If one wraps the $\mathbb{P}^1$ with anti-D5-branes, one should flip the sign of the NS flux in order to have a blown up S^3 (which can be viewed as a generalization of Seiberg duality to $N_f = 0$).

4.4.2 General prescription

Although we focused on the physics around just one $\mathbb{P}^1$ in the above, the result is applicable to general cases where we have multiple $\mathbb{P}^1$ wrapped with D5-branes, because the geometry near each $\mathbb{P}^1$ is identical to the conifold geometry considered above. Therefore, if we replace $\widehat{N}$ with $\widehat{N}_i$, all of the above conclusions carry over.

Once we have understood the physics, we can forget about the D3-brane hypermultiplet $(Q, \widetilde{Q})$ and state the result as a general prescription for how string theory treats $U(0)$, $SO(0)$, $SO(2)$, and $Sp(0), U(1)$ groups in the geometric dual description:

- **$\boldsymbol{U(0)}$, $\boldsymbol{SO(0)}$, $\boldsymbol{SO(2)}$:**
  There are no glueball variables associated with these gauge groups, so we should take the corresponding $S \to 0$.

- **All other groups, including $\boldsymbol{Sp(0)}$ and $\boldsymbol{U(1)}$:**
  We should consider and extremize the corresponding glueball field S.

This prescription should also be applied when using the matrix model [5, 6, 7] to compute the glueball superpotentials.

4.5 Examples

Let us scan over all of the examples of (4.2.1), considering the vacuum where the gauge group is unbroken, and ask when glueball fields S_i for the apparently trivial groups in (4.2.1) should be set to zero, or included and extremized. For the first three cases in (4.2.1), $U(N)$, $SO(N)$, and $Sp(N)$ with adjoint, the breaking (4.2.1) is $G \to G \times U(0)^{K-1}$, and the glueball fields S_i for the $U(0)$ factors are to be set to zero. This justifies the analysis of these theories in the unbroken vacua in [55, 40]. The next case is $SO(N)$ with a symmetric tensor S, where the vacuum with unbroken gauge group is to be understood as $SO(N) \to SO(N) \times SO(0)^{K-1}$, and again the glueball fields S_i for the $SO(0)$ factors are set to zero. This eliminates the Veneziano–Yankielowicz part of the superpotential for $SO(0)$, but the -1 unit of flux associated with each $SO(0)$ does contribute to flux terms $\widehat{N}_i \Pi_i = -\Pi_i$ in (4.3.12), even though this does not contain $SO(0)$ glueballs any more.

The next case is $Sp(N)$ with an antisymmetric tensor, where the vacuum with unbroken gauge group is to be understood as $Sp(N) \to Sp(N) \times Sp(0)^{K-1}$. Unlike the above cases, here we must keep and extremize the S_i for the $Sp(0)$ factors, as will be further discussed shortly.

For the next to last example in (4.2.1), $U(N)$ with $\Phi + S + \widetilde{S}$, the vacuum with unbroken gauge group is to be understood as $U(N) \to SO(0) \times U(N) \times U(0)^{K-1}$, and the glueball fields S_i for $SO(0)$ and $U(0)$ are to be set to zero. Finally, for the last example in (4.2.1), $U(N)$ with $\Phi + A + \widetilde{A}$, the vacuum with unbroken gauge group is to be understood as $U(N) \to Sp(0) \times U(N) \times U(0)^{K-1}$. Though the string engineering of these examples differs somewhat from those discussed in sect. 4 (it was obtained in [65]), the general prescription of sect. 4 is expected to carry over in general: the glueball field S_0 for the $Sp(0)$ factor should be included and extremized, rather than set to zero. On the other hand, the S_i for the $U(0)$

factors are set to zero. These latter two theories in (4.2.1) were considered in [27] and it was noted there that for the case with antisymmetric one expands on the matrix model side around a different vacuum than would be naively expected; this indeed corresponds to keeping and extremizing the glueball field S_0 for the $Sp(0)$ factor, as we have discussed.

We now illustrate some other breaking patterns in the examples of (4.2.1), from the matrix model perspective, for the case of $K = 2$. We also compare with standard gauge theory results and generally find agreement, even in cases where there was room for disagreement because of the possibility of residual instanton effects along the lines of [15]. As will be discussed in more detail in the following section, the agreement is thanks to a remarkable interplay of different residual instanton contributions.

4.5.1 $SO/Sp(N)$ theory with adjoint

Consider $\mathcal{N} = 2$ $SO/Sp(N)$ theory broken to $\mathcal{N} = 1$ by a tree level superpotential for the adjoint chiral superfield Φ:

$$W_{\text{tree}} = \frac{1}{2}\text{Tr}[W(\Phi)], \qquad W(x) = \frac{m}{2}x^2 + \frac{g}{4}x^4. \tag{4.5.1}$$

In the SO case, we can skew-diagonalize Φ as

$$\Phi \sim \text{diag}[\lambda_1, \cdots, \lambda_{N/2}] \otimes i\sigma^2. \tag{4.5.2}$$

The superpotential (4.5.1) has critical points at $\lambda = 0$ and $\lambda = \pm\sqrt{m/g}$. The classical supersymmetric vacuum of the theory is given by distributing N_0 of the N "eigenvalues" λ_i at the critical point $\lambda = 0$ and N_1 "eigenvalue" pairs at $\lambda = \pm\sqrt{m/g}$, with $N_0 + 2N_1 = N$. In this vacuum, the gauge group breaks as $SO(N) \to SO(N_0) \times U(N_1)$. Similarly in the Sp case we have $Sp(N) \to Sp(N_0) \times U(N_1)$.

In the matrix model prescription, the effective superpotential in these vacua is calculated by matrix model as

$$\begin{aligned}W_{\rm DV}(S_0,N_0;S_1,N_1) &= \left(\frac{N_0}{2}\mp 1\right) S_0\left[1-\ln\left(\frac{S_0}{\Lambda_0^3}\right)\right]\\ &\qquad + N_1 S_1\left[1-\ln\left(\frac{S_1}{\Lambda_1^3}\right)\right] + W_{\rm pert},\\ W_{\rm pert} &= N_0\frac{\partial\mathcal{F}_{S^2}}{\partial S_0} + N_1\frac{\partial\mathcal{F}_{S^2}}{\partial S_1} + 4\mathcal{F}_{\mathbb{RP}^2}.\end{aligned} \tag{4.5.3}$$

where $\mathcal{F}_{S^2}$ and $\mathcal{F}_{\mathbb{RP}^2}$ are the S^2 and $\mathbb{RP}^2$ contributions, respectively, to the free energy of the associated $SO/Sp(N)$ matrix model, as defined in Appendix 4.A. The scales Λ_0, Λ_1 in (4.5.3) are the energy scales of the low energy $SO/Sp(N_0)$, $U(N_1)$ theories with the Φ field integrated out, respectively. They are related to the high energy scale Λ by the matching conditions as in (4.2.11), which yields

$$\begin{aligned}(\Lambda_0)^{3(N_0/2\mp 1)} &= m^{N_0/2-N_1\mp 1} g^{N_1}\Lambda^{N\mp 2},\\ (\Lambda_1)^{3N_1} &= 2^{-N_1} m^{-N_0\pm 2} g^{N_0+N_1\mp 2}\Lambda^{2(N\mp 2)}.\end{aligned} \tag{4.5.4}$$

The matrix model free energy is computed in Appendix 4.A, and the result is

$$\begin{aligned}W_{\rm pert} = \left(\frac{N_0}{2}\mp 1\right)&\Bigg[\left(\frac{3}{2}S_0^2 - 8S_0S_1 + 2S_1^2\right)\alpha\\ &+\left(-\frac{9}{2}S_0^3 + 42S_0^2S_1 - 36S_0S_1^2 + 4S_1^3\right)\alpha^2\\ &+\left(\frac{45}{2}S_0^4 - \frac{932}{3}S_0^3S_1 + 523S_0^2S_1^2 - \frac{608}{3}S_0S_1^3 + \frac{40}{3}S_1^4\right)\alpha^3\Bigg]\\ + N_1\Bigg[&\left(-2S_0^2 + 2S_0S_1\right)\alpha + \left(7S_0^3 - 18S_0^2S_1 + 6S_0S_1^2\right)\alpha^2\\ &+\left(-\frac{233}{6}S_0^4 + \frac{524}{3}S_0^3S_1 - 152S_0^2S_1^2 + \frac{80}{3}S_0S_1^3\right)\alpha^3\Bigg]\\ + \mathcal{O}(\alpha^4)\ ,&\end{aligned} \tag{4.5.5}$$

where $\alpha \equiv g/m^2$. The result (4.5.5) agrees with the one obtained in [71], where the glueball superpotential was calculated by evaluating the periods (4.3.6). The

full result, (4.5.3) and (4.5.5), has the expected general form (4.3.12):

$$W_{\text{eff}} = \left(\frac{N_0}{2} \mp 1\right) \Pi_0(S_0, S_1) + N_1 \Pi_1(S_0, S_1) - 2\pi i \alpha \left(\frac{1}{2} S_0 + S_1\right). \qquad (4.5.6)$$

(Here α is the flux defined by (4.3.8), not to be confused with the expansion parameter α in (4.5.5).)

The general prescription in section 4.4 reads in the present case as follows:

- $\boldsymbol{SO/Sp(N) \to SO/Sp(N) \times U(0)}$ **(unbroken** $\boldsymbol{SO/Sp}$**):**
  Set $N_1 = 0$, $S_1 = 0$. Then the superpotential is

  $$W_{\text{eff}}(S_0, N_0) = \left(\frac{N}{2} \mp 1\right) S_0[1 - \ln(S_0/\Lambda_0^3)] + 2\left(\frac{N}{2} \mp 1\right) \left.\frac{\partial \mathcal{F}_{S^2}}{\partial S_0}\right|_{S_1=0}. \qquad (4.5.7)$$

  This superpotential coincides with that of $U(N \mp 2)$ with adjoint and breaking pattern $U(N \pm 2) \to U(N \mp 2) \times U(0) \times U(0)$, as expected from the map (4.3.14) or (4.3.16). As shown in [40] for $SO(N)$, this matrix model result agrees with that of standard gauge theory, via using the corresponding Seiberg–Witten curve.

- $\boldsymbol{SO(N) \to SO(0) \times U(N/2)}$**:**
  Set $N_0 = 0$ and take $S_0 \to 0$, which eliminates the Veneziano–Yankielowicz part for the $SO(0)$. Then the superpotential is

  $$\begin{aligned} W_{\text{eff}}(S_1) &= \frac{N}{2} S_1[1 - \ln(S_1/\Lambda_1^3)] + 4\, \mathcal{F}_{\mathbb{RP}^2}|_{S_0=0}\,. \\ &= -\Pi_0(S_0, S_1)|_{S_0=0} + \frac{N}{2} \Pi_1(S_0, S_1)|_{S_0=0} - 2\pi i \alpha S_1. \qquad (4.5.8) \end{aligned}$$

  Note that $\left.\frac{\partial \mathcal{F}_{S^2}}{\partial S_1}\right|_{S_0=0} = 0$ because $\mathcal{F}_{S^2}$ does not contain terms with S_1 only (all terms are of the form $S_0^n S_1^m$ with $n > 0$). And though the Veneziano–Yankielowicz part of the superpotential for $SO(0)$ is eliminated via $S_0 \to 0$, the -1 units of flux associated with $SO(0)$ does make a contribution in

(4.5.8), with the non-zero terms in $-\Pi_0(S_0, S_1)|_{S_0=0}$ in the second line of (4.5.8) coming from the term $4\mathcal{F}_{\mathbb{RP}^2}|_{S_0=0}$.

- ***SO(N) → SO(2) × U(N/2 − 1):***

  Set $N_0 = 2$, $S_0 = 0$ and remove the Veneziano–Yankielowicz part for the $SO(2)$. Then the superpotential is

$$W_{\rm DV}(S_1) = \left(\frac{N}{2} - 1\right) S_1[1 - \ln(S_1/\Lambda_1^3)], \tag{4.5.9}$$

  where we used $\left.\frac{\partial \mathcal{F}_{S^2}}{\partial S_1}\right|_{S_0=0} = 0$ again. Integrating out S_1 gives

$$W_{\rm low} = \left(\frac{N}{2} - 1\right) \Lambda_1^3 = \frac{1}{2}\left(\frac{N}{2} - 1\right) g\Lambda^4 \tag{4.5.10}$$

- ***Sp(N) → Sp(0) × U(N/2):***

  Set $N_0 = 0$ in the equation and keep both S_0 and S_1. Then the superpotential is

$$\begin{aligned} W_{\rm DV}(S_0, N_0, S_1, N_1) &= S_0[1 - \ln(S_0/\Lambda_0^3)] + NS_1[1 - \ln(S_1/\Lambda_1^3)] + W_{\rm pert}, \\ W_{\rm pert} &= \frac{N}{2}\frac{\partial \mathcal{F}_{S^2}}{\partial S_1} + 4\mathcal{F}_{\mathbb{RP}^2} = \frac{N}{2}\frac{\partial \mathcal{F}_{S^2}}{\partial S_1} + 2\frac{\partial \mathcal{F}_{S^2}}{\partial S_0}. \end{aligned} \tag{4.5.11}$$

For various breaking patterns, we integrated out the glueball superfield(s) from the glueball superpotential (4.5.7)–(4.5.11), and calculated the low energy superpotential $W_{\rm low}$ as a function of coupling constants m, g, and the scale Λ. Having obtained the actual matrix model results, we can compare to the superpotential as computed via standard gauge theory methods, such as via factorizing of the Seiberg–Witten curve. This method is reviewed in Appendix 4.B, and the results are found to agree with the matrix model results completely. The resulting $W_{\rm low}$ is shown in Table 4.1.

breaking pattern	W'_{low}
$SO(4) \to SO(4) \times U(0)$	$m\Lambda^2 + \frac{3}{2} g\Lambda^4$
$SO(4) \to SO(2) \times U(1)$	$\frac{1}{2} g\Lambda^4$
$SO(4) \to SO(0) \times U(2)$	$m\Lambda^2 - \frac{1}{2} g\Lambda^4$
$SO(6) \to SO(6) \times U(0)$	$2m\Lambda^2 + 3g\Lambda^4$
$SO(6) \to SO(4) \times U(1)$	$g\Lambda^4$
$SO(6) \to SO(2) \times U(2)$	$g\Lambda^4$
$SO(6) \to SO(0) \times U(3)$	$\frac{3}{2}(m^2 g\Lambda^8)^{1/3} - \frac{1}{2}(g^5\Lambda^{16}/m^2)^{1/3} - g^3\Lambda^8/6m^2 + \cdots$
$SO(8) \to SO(8) \times U(0)$	$3m\Lambda^2 + \frac{9}{2} g\Lambda^4$
$SO(8) \to SO(6) \times U(1)$	$2(mg)^{1/2}\Lambda^3 + g^2\Lambda^6/m - (g^7/m^5)^{1/2}\Lambda^9 + \cdots$
$SO(8) \to SO(4) \times U(2)$	$2g^2\Lambda^6/m - 4g^5\Lambda^{12}/m^4 + 32g^8\Lambda^{18}/m^7 + \cdots$
$SO(8) \to SO(2) \times U(3)$	$\frac{3}{2} g\Lambda^4$
$SO(8) \to SO(0) \times U(4)$	$2(mg)^{1/2}\Lambda^3 - g^2\Lambda^6/2m + \frac{1}{4}(g^7/m^5)^{1/2}\Lambda^9 + \cdots$
$Sp(2) \to Sp(2) \times U(0)$	$2m\Lambda^2 + 3g\Lambda^4$
$Sp(2) \to Sp(0) \times U(1)$	$g\Lambda^4$
$Sp(4) \to Sp(4) \times U(0)$	$3m\Lambda^2 + \frac{9}{2} g\Lambda^4$
$Sp(4) \to Sp(2) \times U(1)$	$2(mg)^{1/2}\Lambda^3 + g^2\Lambda^6/m - (g^7/m^5)^{1/2}\Lambda^9 + \cdots$
$Sp(4) \to Sp(0) \times U(2)$	$2g^2\Lambda^6/m - 4g^5\Lambda^{12}/m^4 + 32g^8\Lambda^{18}/m^7 + \cdots$

Table 4.1: The low energy superpotential calculated from the factorization of the Seiberg–Witten curve and from matrix model. In the above, the classical contribution has been subtracted: $W_{\text{low}} = -N_1 m^2/4g + W'_{\text{low}}$.

We have been considering $SO(N)$ theory with even N. The $SO(2N+1)$ theory with adjoint in the $SO(2N+1) \to U(1)^N$ vacuum was studied diagrammatically in [72].

4.5.2 $Sp(N)$ theory with antisymmetric tensor

Consider $Sp(N)$ theory with an antisymmetric tensor chiral superfield $A = -A^T$. Take cubic tree level superpotential

$$W_{\text{tree}} = \frac{1}{2}\text{Tr}[W(\Phi)], \qquad W(x) = \frac{m}{2}x^2 + \frac{g}{3}x^3, \tag{4.5.12}$$

where $\Phi = AJ$, and J is the invariant antisymmetric tensor $J = \mathbb{1}_{N/2} \otimes i\sigma^2$. We do not require A to be traceless, i.e. $\text{Tr}[\Phi] = \text{Tr}[AJ] \neq 0$. By a complexified $Sp(N)$ gauge rotation, Φ can be diagonalized as [21]

$$\Phi \cong \text{diag}[\lambda_1, \cdots, \lambda_{N/2}] \otimes \mathbb{1}_2, \qquad \lambda_i \in \mathbb{C}. \tag{4.5.13}$$

The superpotential (4.5.12) has critical points at $\lambda = 0, -m/g$. The classical supersymmetric vacuum of the theory is given by distributing N_1 and N_2 eigenvalues at the critical point $\lambda = 0$ and $\lambda = -m/g$, respectively, with $N_1 + N_2 = N$, breaking $Sp(N) \to Sp(N_1) \times Sp(N_2)$.

The glueball superpotential is calculated from the associated $Sp(N)$ matrix model as

$$\begin{aligned} W_{\text{DV}}(S_1, N_1; S_2, N_2) = \left(\frac{N_1}{2} + 1\right) S_1[1 - \ln(S_1/\Lambda_1^3)] \\ + \left(\frac{N_2}{2} + 1\right) S_2[1 - \ln(S_2/\Lambda_2^3)] + W_{\text{pert}}, \\ W_{\text{pert}} = N_1 \frac{\partial \mathcal{F}_{S^2}}{\partial S_1} + N_2 \frac{\partial \mathcal{F}_{S^2}}{\partial S_2} + 4\mathcal{F}_{\mathbb{RP}^2}. \end{aligned} \tag{4.5.14}$$

The scales Λ_1, Λ_2 in (4.5.3) respectively are the energy scales of the low energy $Sp(N_1)$, $Sp(N_2)$ theories with the Φ field integrated out. They are related to the

high energy scale Λ by the matching conditions as in (4.2.13), which yields

$$\begin{aligned}
(\Lambda_1)^{3(N_1/2+1)} &= m^{N_1/2-N_2-1} g^{N_2} \Lambda^{N+4}, \\
(\Lambda_2)^{3(N_2/2+1)} &= (-1)^{N_2/2-1} m^{-N_1+N_2/2-1} g^{N_1} \Lambda^{N+4}.
\end{aligned} \tag{4.5.15}$$

The matrix model free energy is computed in Appendix 4.A, and the result is

$$\begin{aligned}
W_{\text{pert}} = (N_1+2)\Bigg[&\left(-S_1^2 + 5S_1S_2 - \frac{5}{2}S_2^2\right)\alpha \\
&+ \left(-\frac{16}{3}S_1^3 + \frac{91}{2}S_1^2S_2 - 59S_1S_2^2 + \frac{91}{6}S_2^3\right)\alpha^2 \\
&+ \left(-\frac{140}{3}S_1^4 + \frac{1742}{3}S_1^3S_2 - 1318S_1^2S_2^2 \right. \\
&\qquad \left. + \frac{2636}{3}S_1S_2^3 - \frac{871}{6}S_2^4\right)\alpha^3\Bigg] \\
+ (N_2+2)\Big[&S_1 \leftrightarrow S_2,\ \alpha \to -\alpha\Big],
\end{aligned} \tag{4.5.16}$$

where $\alpha \equiv g^2/m^3$. This is as expected from the map (4.3.17) and [24].

In particular, let us concentrate on the unbroken case, $N_2 = 0$. Unlike [23, 31], we do not set $S_2 = 0$, but rather keep S_2 non-zero and extremize with respect to it, according to our general prescription, to obtain the actual matrix model result. After integrating out S_1 and S_2 from (4.5.16), we obtain the superpotential as a power series in Λ_1 and Λ_2 as

$$W_{\text{low}} = \left(\frac{N}{2}+1\right)\Lambda_1^3 + \Lambda_2^3 + (\text{higher order terms in } \Lambda_{1,2}). \tag{4.5.17}$$

The matching relation (4.5.15) gives

$$\Lambda_2^3 = -(\Lambda_1^3)^{N/2+1}\alpha^{N/2}, \tag{4.5.18}$$

so the terms containing Λ_2 in (4.5.17) starts to contribute to the superpotential at order $(\Lambda_1^3)^{N/2+1}$, i.e. like $Sp(N)$ instantons. If we use the relation (4.5.18) and

write out all the terms in (4.5.17), we obtain

$$
\begin{aligned}
N=0: &\quad W_{\text{low}} = \mathcal{O}(\alpha^4), \\
N=2: &\quad W_{\text{low}} = 2\Lambda_1^3 + \mathcal{O}(\alpha^4), \\
N=4: &\quad W_{\text{low}} = 3\Lambda_1^3 - \Lambda_1^6\alpha - 2\Lambda_1^9\alpha^2 - \frac{187}{27}\Lambda_1^{12}\alpha^3 + \mathcal{O}(\alpha^4), \\
N=6: &\quad W_{\text{low}} = 4\Lambda_1^3 - 3\Lambda_1^6\alpha - \frac{47}{6}\Lambda_1^9\alpha^2 - \frac{75}{2}\Lambda_1^{12}\alpha^3 + \mathcal{O}(\alpha^4), \\
N=8: &\quad W_{\text{low}} = 5\Lambda_1^3 - 5\Lambda_1^6\alpha - 13\Lambda_1^9\alpha^2 - 65\Lambda_1^{12}\alpha^3 + \mathcal{O}(\alpha^4).
\end{aligned}
\tag{4.5.19}
$$

Thus properly accounting for S_2, it turns out that these matrix model results agree perfectly, up to the order presented, with the standard gauge theory results (Eq. (4.13) of [23]). (The discrepancies found in [23] set in at order $\Lambda_1^{3(N/2+1)}$, and are canceled e.g. by (4.5.18).) In (4.5.19), for $N = 0, 2$, there were rather remarkable cancellations between the instanton contributions from $Sp(N_1)$ and $Sp(N_2)$. This will be further discussed and generalized in the next section.

The matrix model prediction for the superpotential of the $SO(N)$ theory with symmetric tensor can similarly be obtained by simply changing the $N_i/2+1$ in (4.5.16) to $N_i/2-1$. It should be possible to compute the superpotential from gauge theory using the duality for this theory [73]. The result is expected to be compatible with the map (4.3.20) to the superpotential computed for the $U(N-2K)$ theory with adjoint.

4.6 Residual instantons: string theory (matrix model) versus gauge theory

A remarkable aspect of the string theory (matrix model) computation of the effective superpotential is that (4.1.3) can be obtained purely in terms of the dynamics of the low-energy $\prod_{i=1}^{K} G(N_i)$ theory on the RHS of (4.1.1). The only

information needed about the high-energy $G(N)$ gauge theory is the perturbative contribution of the $G(N)/\prod_i G_i(N_i)$ ghosts to the glueball superpotential (4.1.3), as discussed in [74], along with the matching relations connecting the scales Λ_i of the low-energy $G_i(N_i)$ factors to the scale Λ of the high-energy $G(N)$ theory. This is very different from the conventional description of standard gauge theory, where there can be non-perturbative contributions to W_{low} which are not readily seen in terms of the low-energy theory on the RHS of (4.1.1). An example of such an effect is instantons in the broken part of the group when $\pi_3(G(N)/\prod_i G_i(N_i)) \neq 0$ (see e.g. [75]). Nevertheless, the string theory/ matrix model does properly reproduce such effects, via a low-energy description.

A gauge theory interpretation for the string theory/ matrix model results was given in [57, 15]: the string theory / matrix model results actually refer to a particularly natural UV completion of the original $G(N)$ theory, where it is embedded in the supergroup $G(N+k|k)$ with k large. This latter theory has a Higgs branch, where k can be reduced successively, eventually Higgsing the theory down to the original $G(N)$ theory. More generally, the theory with breaking pattern (4.1.1) is replaced with

$$G(N+k|k) \to \prod_{i=1}^{K} G_i(N_i+k_i|k_i), \tag{4.6.1}$$

which has a Higgs branch flat direction connecting it to (4.1.1). Consideration of the particular matter content of the $G(N+k|k)$ theories along the Higgs branch, which often has extended supersymmetry, suggests that no dynamically generated superpotential ever lifts this Higgs branch moduli space, i.e. that the superpotentials of these particular theories are always independent of the location of the theory on this Higgs branch [15]. Moving along the Higgs branch has the effect of reducing k, and this expected independence of the superpotential of the position on the Higgs branch fits with the fact that the $G(N+k|k)$ matrix model

results are k independent, because all k dependence cancels in the supertraces.

Because of the expected independence of the superpotential on the Higgs branch, and because we Higgs back to the original $G(N)$ theory, in most cases, this "F-completion" of the original $G(N)$ theory into the $G(N+k|k)$ theory is of no consequence. There are, however, a few rare exceptions, where the superpotential of the Higgsed $G(N+k|k)$ theory differs from that of the standard $G(N)$ theory. This difference comes from residual instantons in $G(N+k|k)/G(N)$, which need not decouple even if $G(N+k|k)$ is Higgsed to $G(N)$ far in the UV. As verified in [15], these residual instanton contributions precisely account for the few differences between the string theory (matrix model) results and standard gauge theory, for example the glueball superpotentials, with coefficient $h = 1$, for $U(1)$ and $Sp(0)$, e.g. with an adjoint and quadratic superpotential.

In many cases, however, these residual instanton contributions sum up to yield precisely the result expected from standard gauge theory, including superpotential contributions which in standard gauge theory would not have had a known low-energy description. In particular, residual instanton contributions which could have lead to potential discrepancies with standard gauge theory often completely cancel. The cancellation occurs once one sums over the different terms i in (4.1.2), upon using the precise matching relation between the low-energy scales Λ_i, and the original high-energy scale Λ.

As an example, consider $U(K)$ with adjoint matter and breaking pattern $U(K) \to U(1)^K$. For n$K = 1$ the string theory (matrix model) description includes a residual instanton effect, yielding $W_{\text{low}} = \Lambda_L^3$ rather than the standard gauge theory answer $W_{\text{sgt}} = 0$ [15]. But for all $K > 1$ the string theory/ matrix model result is $W_{\text{low}} = 0$, in agreement with the standard gauge theory expectation for $U(K) \to U(1)^K$. The result $W_{\text{low}} = 0$ looks like a remarkable can-

cellation because the glueball superpotential $W_{\rm eff}(S_1, \dots S_K)$ is quite non-trivial. Nevertheless, upon solving for the $\langle S_i\rangle$ and plugging back in, the exact result for $W_{\rm low} = W_{\rm eff}(\langle S_i\rangle)$ is zero, as was proven in [76].

To illustrate this cancellation, consider the leading order gaugino condensation contribution to $W_{\rm eff}(S_i)$ in the string theory (matrix model) constructions, where the unbroken $U(1)$ factors in $U(K) \to U(1)^K$ contribute as in (4.2.14), with $h_i = 1$, unlike in standard gauge theory:

$$W_{\rm gc}(S_i) = \sum_{i=1}^{k} S_i \left(\log(\frac{\Lambda_i^3}{S_i}) + 1\right), \tag{4.6.2}$$

with $\Lambda_i^3 = g_{K+1}\Lambda^{2N}\prod_{j\neq i}(a_j - a_i)^{-1}$ by using (4.2.10) with all $N_i = 1$. Though this is a non-trivial superpotential, it vanishes upon integrating out the S_i:

$$W_{\rm gc}(\langle S_i\rangle) = \sum_{i=1}^{K}\Lambda_i^3 = g_{K+1}\Lambda^{2N}\sum_{i=1}^{K}\prod_{j\neq i}(a_j - a_i)^{-1} = g_{k+1}^2\Lambda^{2N}\oint \frac{dx}{2\pi i}\frac{1}{W'(x)} = 0. \tag{4.6.3}$$

The contour in (4.6.3) encloses all the zeros of $W'(x)$, and we get zero for all $K > 1$ by pulling the contour off to infinity. We see here why $K = 1$ is different: we then get a residue at infinity, leading to the low energy superpotential $W_{U(1)} = \Lambda_L^3$, as in (4.2.9), with $h_{U(1)} = 1$ as in (4.1.4).

To give another example of such a cancellation of residual instanton effects, consider the string theory (matrix model) result for $Sp(N)$ with an antisymmetric tensor A, with $W_{\rm tree}$ having K critical points, for the case $N = 0$. For the case of $K = 1$, the superpotential is just a mass term for A and the low-energy superpotential is the $Sp(0)$ gaugino condensation superpotential, with $h(Sp(0)) = 1$: $W = \Lambda^3$, unlike standard gauge theory. Again, this can be understood as a residual instanton effect in the F-completion of $Sp(N)$ to $Sp(N+k|k)$, which is present precisely for the case $N = 0$ [15]. For a higher order superpotential, $K > 1$, we

would write the breaking pattern as $Sp(0) \to Sp(0)^K$. For all $K > 1$, the residual instanton effects all cancel, precisely as in the $U(K) \to U(1)^K$ example discussed above; in fact, the two theories have the same effective superpotential $W_{\rm eff}$ (aside from the classical difference), as discussed in [24] and sect. 3.3. Thus, for example, (4.6.3) can also be interpreted as the leading gaugino condensation contributions from the $Sp(0)^K$ factors, and where we now use the matching relation (4.2.13) to relate the Λ_i^3 to $g_{K+1}\Lambda^{2K}\prod_{j\neq i}(a_j - a_i)^{-1}$. Again, there is complete cancellation in $W_{\rm eff}$ here, except for the case $K = 1$.

More generally, for $Sp(N)$ with antisymmetric, breaking as $Sp(N) \to \prod_{i=1}^K Sp(N_i)$, the results obtained via the string theory / matrix model glueball potential $W_{\rm eff}(S_1, \ldots S_K)$, upon integrating out the S_i, appears to always agree with standard gauge theory results for the superpotential [21, 22], as seen in the examples of [24] and (4.5.19). This agreement comes about via a remarkable interplay between the different terms i in (4.3.12). If we treated the scales Λ_i of the $Sp(N_i)$ factors as if they were initially independent, each term $\widehat{N}_i\Pi(\langle S_i\rangle) - 2\pi i\eta_i\langle S_i\rangle$ in (4.3.12) would be a complicated function of Λ_i, which does not have a known, conventional, interpretation in terms of standard gauge theory for the low-energy $Sp(N_i)$ factor. But upon adding the different i terms and using the matching relations relating Λ_i to Λ, e.g. (4.2.13), one nevertheless obtains the standard gauge theory results, thanks to an intricate interplay between the different terms i.

By the map of (4.3.17) [24], the agreement between string theory / matrix models and standard gauge theory for $Sp(N)$ with antisymmetric can be phrased as such an agreement for $U(N/2+K)$ with adjoint and breaking pattern $U(N/2+K) \to \prod_{i=1}^K U(N_i/2+1)$.

As another example, consider $U(N)$ with adjoint Φ and superpotential having

$K = N-1$, in the vacuum where $U(N) \to U(2) \times U(1)^{N-2}$. Factorizing the Seiberg–Witten curve yields for the exact superpotential [77]

$$W_{\text{exact}} = W_{\text{cl}}(g) \pm 2g_N \Lambda^N. \tag{4.6.4}$$

The map of [24] and subsection 4.3.3 relates this to $Sp(2) \to Sp(2) \times Sp(0)^{N-2}$, where the exact gauge theory result agrees with (4.6.4), up to the classical shift, upon using the relation (4.3.18). A priori, one might expect the string theory / matrix model result to disagree with (4.6.4), due to residual instanton contributions from the $U(1)^{N-2}$ or the $Sp(0)^{N-2}$ in $U(N) \to U(2) \times U(1)^{N-2}$ and $Sp(2) \to Sp(2) \times Sp(0)^{N-2}$ respectively. But the string theory / matrix model result nevertheless agrees with (4.6.4), thanks to the interplay between the different terms. Consider, in particular, the case $U(3) \to U(2) \times U(1)$. The fact that (4.6.4) will only hold if remarkable cancellations occur upon integrating out S_1 and S_2 from the non-trivial $W(S_1, S_2)$, was discussed in [10], where the cancellations were verified to indeed occur, up to order α^3. This is checked to one higher order in (4.5.19), since it is related to $Sp(2) \to Sp(2) \times Sp(0)$ by the map of [24] and subsection 4.3.3. The leading order cancellation, say in terms of $U(3) \to U(2) \times U(1)$, is between $U(1)$ gaugino condensation, Λ_2^3, and a higher order term coming from integrating out S_1 from $W_{\text{pert}}(S_i)$.

The residual instanton contributions associated with the UV completion (4.6.1), as opposed to the standard gauge theory results for (4.1.1) do not always cancel, however. The cases where we find non-cancellations are when the degree of the superpotential is sufficiently large, so that it contains terms which are not independent moduli. As an example, consider $U(1)$ with W_{tree} as in (4.2.2) having K minima, breaking $U(1) \to U(1) \times U(0)^{K-1}$. The gaugino condensation contribution to the superpotential, according to the string theory (matrix model) construction, is given by (4.2.14) with $h_1 = 1$ and all other $h_i = 0$ and their S_i

set to zero. Upon integrating out S_1, we thus obtain the superpotential

$$W_{\rm gc} = \Lambda_1^3 = \Lambda^2 W''(a_1) = g_{K+1}\Lambda^2 \prod_{j\neq 1}(a_j - a_1), \tag{4.6.5}$$

where we used the matching relation (4.2.10) with $N_1 = 1$ and all other $N_j = 0$.

The full low-energy effective superpotential $W_{\rm low}(g_i, \Lambda)$ can be regarded as the generating function for the operator expectation values:

$$\langle u_j\rangle = \frac{\partial W_{\rm low}(g_i,\Lambda)}{\partial g_j} \qquad u_i \equiv \frac{1}{j}{\rm Tr}\Phi^j. \tag{4.6.6}$$

In the $U(1)$ theory, we have classical relations $u_j = \frac{1}{j}u_1^j$. But the quantum contribution (4.6.5) (along with additional, higher order contributions) imply quantum deformation of these classical relations, due to the residual instanton effects in the $U(1+k|k) \to U(1+k_1|k_1) \times U(k_2|k_2)\ldots U(k_K|k_K)$ F-completion. For the simplest such example, consider $U(1)$ with $W_{\rm tree} = \frac{1}{2}m\Phi^2 + \lambda\Phi$. The low-energy superpotential is

$$W_{\rm low} = -\frac{\lambda^2}{2m} + m\Lambda^2, \tag{4.6.7}$$

with the first term the classical contribution and the second the residual instanton. Using (4.6.6) we then get

$$\langle u_1\rangle = -\frac{\lambda}{m}, \quad , \langle u_2\rangle = \frac{\lambda^2}{2m^2} + \Lambda^2 \qquad \text{i.e.} \quad \langle u_2\rangle = \frac{1}{2}\langle u_1^2\rangle + \Lambda^2, \tag{4.6.8}$$

which can be regarded as an instanton correction to the composite operator u_2.

As another such example, consider $U(2)$ with an adjoint and $W_{\rm tree}$ having K minima, in the vacuum where the gauge group is broken as $U(2) \to U(1)\times U(1)\times U(0)^{K-2}$. The gaugino condensation contribution to $W_{\rm low}$ is

$$W_{\rm gc} = \Lambda_1^3 + \Lambda_2^3 = g_{K+1}\Lambda^4 \left(\frac{\prod_{j=3}^K (a_j - a_1) - \prod_{j=3}^K (a_j - a_2)}{a_2 - a_1}\right). \tag{4.6.9}$$

For example, for $U(2)$ with W_{tree} having $K=3$ critical points, we break $U(2) \to U(1) \times U(1) \times U(0)$ and (4.6.9) leads to

$$W_{\text{gc}} = g_4 \Lambda^4. \tag{4.6.10}$$

Computing expectation values as in (4.6.6) this leads to

$$\langle u_4 \rangle = \langle u_4 \rangle_{\text{cl}} + \Lambda^4, \tag{4.6.11}$$

which can be interpreted as an instanton contribution to the composite operator $u_4 = \text{Tr}\Phi^4$ in $U(2)$ gauge theory. More generally, for $U(N)$ gauge theory, the independent basis of operators $u_j = \frac{1}{j}\text{Tr}\Phi^j$ are only those with $j \leq N$, those with $j > N$ can be expressed as products of these basis operators via classical relations. But these relations can be affected by instantons. In particular, for $U(N)$ with an adjoint, the instanton factor is Λ^{2N}, so operators u_j with $j \geq 2N$ can be affected. The above residual instanton contributions of the $U(N+k|k)$ UV completion can be interpreted as implying specific such instanton corrections to the higher Casimirs u_j.

A similar situation arises in the $\mathcal{N} = 1^*$ $U(N)$ theory, where the effective superpotential of the matrix model and conventional gauge theory differ by a contribution $N^2 m^3 E_2(N\tau)$ [58]; this was interpreted in [58] as differing operator definitions of $\text{Tr}\Phi^2$ between gauge theory and the matrix model at the level of instantons. Related issues for multi-trace operators were seen in [41].

4.7 Conclusions

To compute the correct string theory / matrix model results, we should include or not include the glueball fields S_i according to the prescription of this paper. Upon doing so, in all examples that we know of, the string theory / matrix

model results agree with the results of standard gauge theory, at least in those cases where the relevant gauge theory does not suffer from UV ambiguities. In the case where such ambiguities are present, for example in defining composite operators appearing in W_{tree}, the string theory / matrix model results correspond to a particular UV definition of the theory. The agreement with standard gauge theory results is often due to a remarkable interplay between the different low-energy terms, found upon integrating out the glueball fields S_i, and connecting their scales Λ_i via the appropriate matching relation to the scale Λ of the original theory. In some cases, this interplay leads to complete cancellations of the residual instanton contributions to W_{low} coming from the $G(N+k|k)$ completion [57, 15]. Perhaps there is some additional structure governing the glueball superpotentials, which would make these remarkable cancellations more manifest.

Appendix

4.A Matrix model calculation of superpotential

In this appendix, after giving a proof for a general relation that relates S^2 and $\mathbb{RP}^2$ contributions to the SO/Sp matrix model free energy, we compute explicitly the free energy of the matrix models associated with SO/Sp gauge theory with adjoint and Sp gauge theory with antisymmetric tensor. These matrix model results are used in section 4.5 to evaluate the glueball superpotential of the corresponding gauge theories.

4.A.1 Proof for relation between $\mathcal{F}_{S^2}$ and $\mathcal{F}_{\mathbb{RP}^2}$

Here we prove a general relation between the S^2 and $\mathbb{RP}^2$ contributions to the $SO/Sp(N)$ matrix model free energy:

$$\mathcal{F}_{\mathbb{RP}^2} = \begin{cases} \mp\dfrac{1}{2}\dfrac{\partial \mathcal{F}_{S^2}}{\partial S_0} & SO/Sp \text{ with adjoint,} \\ \mp\dfrac{1}{2}\displaystyle\sum_{i=1}^{K}\dfrac{\partial \mathcal{F}_{S^2}}{\partial S_i} & SO/Sp \text{ with symmetric/antisymmetric tensor.} \end{cases} \tag{4.A.1}$$

The first equation was conjectured in [14, 39] based on explicit diagrammatic calculations, and proven in [40] for the case of unbroken vacua. Here we will give a general matrix model proof for arbitrary breaking pattern. These relations are equivalent to the maps (4.3.14), (4.3.16), (4.3.17), and (4.3.20), which we obtained in subsect. 4.3.3 immediately from the string theory geometric transition construction, accounting for the orientifold contributions to the fluxes.

Consider $U(\mathbf{N})$ and $SO/Sp(\mathbf{N})$ matrix models which correspond to $U(N)$ and $SO/Sp(N)$ gauge theories with a two-index tensor matter field. The partition function is

$$\mathbf{Z} = e^{-\frac{1}{\mathbf{g}^2}\mathbf{F}(\mathbf{S}_i)} = \int d\mathbf{\Phi}\, e^{-\frac{1}{\mathbf{g}}W_{\text{tree}}(\mathbf{\Phi})}. \tag{4.A.2}$$

We denote matrix model quantities by boldface letters, following the notation of [31]. $\mathbf{\Phi}$ is an $\mathbf{N}\times\mathbf{N}$ matrix corresponding to the Φ field in gauge theory, and the "action" W_{tree} is defined in (4.2.2), (4.2.4) and (4.2.5). The matrix integral (4.A.2) is evaluated perturbatively around the general broken vacua of (4.2.1), with N_i replaced by $\mathbf{N}_i$. We take the large $\mathbf{N}$ limit $\mathbf{N}_i \to \infty$, $\mathbf{g} \to 0$ with $\mathbf{g}\mathbf{N}_i \equiv \mathbf{S}_i$ kept finite. The dependence of the free energy $\mathbf{F}(\mathbf{S}_i)$ on $\mathbf{N}_i$ are eliminated in favor

of $\mathbf{S}_i$, and $\mathbf{F}(\mathbf{S}_i)$ is expanded in the 't Hooft expansion as

$$\mathbf{F}(\mathbf{S}_i) = \sum_{\mathcal{M}} \mathbf{g}^{2-\chi(\mathcal{M})} \mathcal{F}_{\mathcal{M}}(\mathbf{S}_i) = \mathcal{F}_{S^2} + \mathbf{g}\mathcal{F}_{\mathbb{RP}^2} + \cdots \tag{4.A.3}$$

where the sum is over all compact topologies $\mathcal{M}$ of the matrix model diagrams written in the 't Hooft double-line notation, and $\chi(\mathcal{M})$ is the Euler number of $\mathcal{M}$.

The matrix model resolvent is defined as follows:

$$\mathbf{R}(z) \equiv \mathbf{g}\left\langle \mathrm{Tr}\left[\frac{1}{z-\mathbf{\Phi}}\right]\right\rangle = \mathbf{R}_{S^2}(z) + \mathbf{g}\mathbf{R}_{\mathbb{RP}^2}(z) + \cdots . \tag{4.A.4}$$

For $U(N)$ theory with adjoint, the expansion parameter is $\mathbf{g}^2$ instead of $\mathbf{g}$, and in particular, $\mathbf{R}_{\mathbb{RP}^2}(z) \equiv 0$. The resolvent and the free energy are related as

$$\mathbf{R}_{\mathcal{M}}(z) = \frac{S}{z}\delta_{\chi(\mathcal{M}),2} + \frac{1}{z^2}\frac{\partial \mathcal{F}_{\mathcal{M}}}{\partial g_1} + \frac{2}{z^3}\frac{\partial \mathcal{F}_{\mathcal{M}}}{\partial g_2} + \frac{3}{z^4}\frac{\partial \mathcal{F}_{\mathcal{M}}}{\partial g_3} + \cdots , \tag{4.A.5}$$

where $S = \sum_{i=1}^{K} S_i$. The resolvents can be determined uniquely by solving the matrix model loop equations (the loop equations for the relevant matrix models are summarized in [31]), under the condition

$$\oint_{A_i} \frac{dz}{2\pi i}\mathbf{R}_{S^2}(z) = \mathbf{S}_i, \qquad \oint_{A_i} \frac{dz}{2\pi i}\mathbf{R}_{\mathbb{RP}^2}(z) = 0. \tag{4.A.6}$$

A_i is the contour around the i-th critical point of $W(z)$. In general, $\mathbf{R}(z)$ develops a cut around each critical point in the large $\mathbf{N}_i$ limit, and A_i is taken to encircle the i-th cut. Note that the expression (4.A.5) should be understood as a Laurent expansion around $z = \infty$, and converges only if $|z|$ is larger than r such that all the singularities (cuts) of the resolvent are inside the circle $C: \ |z| = r$.

On the other hand, gauge theory resolvents $R(z)$, $T(z)$ (see e.g. [17]) are determined uniquely by solving the Konishi anomaly equations (the Konishi anomaly equations for the relevant gauge theories are summarized in [31]), under the condition

$$\oint_{A_i} \frac{dz}{2\pi i} R(z) = S_i, \qquad \oint_{A_i} \frac{dz}{2\pi i} T(z) = N_i. \tag{4.A.7}$$

As was shown in [31], the matrix model resolvents $\mathbf{R}_{S^2}(z)$, $\mathbf{R}_{\mathbb{RP}^2}(z)$ are related to the gauge theory resolvents $R(z)$, $T(z)$ as

$$R(z) = \mathbf{R}_{S^2}(z), \qquad T(z) = \sum_{U,\,SO,\,Sp} N_i \frac{\partial}{\partial S_i}\mathbf{R}_{S^2}(z) + 4\mathbf{R}_{\mathbb{RP}^2}(z) \tag{4.A.8}$$

with S_i and $\mathbf{S}_i$ identified[6].

First, consider $SO(2N)/Sp(N)$ theory with adjoint. The general breaking pattern is $SO/Sp(2N) \to SO/Sp(2N_0) \times U(N_1) \times \cdots \times U(N_K)$ (Eq. (4.2.1)), where $N = N_0 + 2\sum_{i=1}^{K} N_i$. Note that the eigenvalues are distributed in a symmetric manner under $z \leftrightarrow -z$, and hence (4.A.7) is

$$\begin{aligned}
&\oint_{A_0} \frac{dz}{2\pi i} R(z) = S_0, \qquad &&\oint_{A_i} \frac{dz}{2\pi i} R(z) = \oint_{A_{-i}} \frac{dz}{2\pi i} R(z) = S_i, \\
&\oint_{A_0} \frac{dz}{2\pi i} T(z) = N_0, \qquad &&\oint_{A_i} \frac{dz}{2\pi i} T(z) = \oint_{A_{-i}} \frac{dz}{2\pi i} T(z) = N_i,
\end{aligned} \tag{4.A.9}$$

where $i = 1, \ldots, K$. The contours A_i and A_{-i} encircle counterclockwise the cuts around $z = a_i$ and $z = -a_i$, respectively. The relation (4.A.8) holds as it is, with the summation understood as over $SO/Sp(N_0)$ and $U(N_i)$, $i = 1, \ldots, K$.

It was shown in [29] that the resolvents of this $SO/Sp(N)$ theory are related to the resolvents $\tilde{R}(z)$ and $\widetilde{T}(z)$ of $U(\widetilde{N} \equiv N \mp 2)$ theory with adjoint as follows:

$$R(z) = \tilde{R}(z), \qquad T(z) = \widetilde{T}(z) \pm \frac{2}{z}. \tag{4.A.10}$$

The tree level superpotential of the $U(\widetilde{N})$ theory is related to the one for the $SO/Sp(N)$ theory as $W^U(z) = W^{SO/Sp}(z)$ (see (4.2.2) and (4.2.5)), and the breaking pattern is $U(\widetilde{N}) \to U(N_{-K}) \times \cdots \times U(N_{-1}) \times U(N_0 \mp 2) \times U(N_1) \times \cdots \times U(N_K)$ with $N_{-i} = N_i$, $1 \le i \le K$. Note that since there is no $z \leftrightarrow -z$ symmetry

[6] The relation (4.A.8) is an obvious generalization of the formula in [31], which was for unbroken vacua, to an arbitrary breaking pattern. The gauge theory resolvents $R(z)$, $T(z)$ given in (4.A.8) clearly satisfy the condition (4.A.7) provided that the matrix model resolvents $\mathbf{R}_{S^2}(z)$, $\mathbf{R}_{\mathbb{RP}^2}(z)$ satisfy the condition (4.A.6).

in the $U(\widetilde{N})$ theory, the $U(N_{-i})$ factors that are "images" for $SO/Sp(N)$ are "real" for $U(\widetilde{N})$. In addition, the glueball $\widetilde{S}$ of the $U(\widetilde{N})$ theory is related to the glueball S of the $SO/Sp(N)$ theory as $\widetilde{S}_0 = S_0$, $\widetilde{S}_i = \widetilde{S}_{-i} = S_i$, $1 \le i \le K$. Therefore, e.g. the first equation in (4.A.10) is more precisely

$$R(z, S_j) = \widetilde{R}(z, \widetilde{S}_j)\big|_{\widetilde{S}_0=S_0,\, \widetilde{S}_i=\widetilde{S}_{-i}=S_i}\,. \tag{4.A.11}$$

Differentiating (4.A.11) with respect to S_j, we obtain

$$\frac{\partial R}{\partial S_0} = \frac{\partial \widetilde{R}}{\partial \widetilde{S}_0}\bigg|_{\substack{\widetilde{S}_0=S_0,\\ \widetilde{S}_i=\widetilde{S}_{-i}=S_i}}, \qquad \frac{\partial R}{\partial S_j} = \left(\frac{\partial \widetilde{R}}{\partial \widetilde{S}_j} + \frac{\partial \widetilde{R}}{\partial \widetilde{S}_{-j}}\right)\bigg|_{\substack{\widetilde{S}_0=S_0,\\ \widetilde{S}_i=\widetilde{S}_{-i}=S_i}}, \tag{4.A.12}$$

where $1 \le j \le K$.

Now, using (4.A.8), let us translate the relation (4.A.10) among gauge theory resolvents into a relation among matrix model resolvents:

$$\mathbf{R}_{S^2} = \tilde{\mathbf{R}}_{S^2}, \tag{4.A.13}$$

$$\begin{aligned} N_0\frac{\partial \mathbf{R}_{S^2}}{\partial S_0} + \sum_{i=1}^{K} N_i \frac{\partial \mathbf{R}_{S^2}}{\partial S_i} + 4\mathbf{R}_{\mathbb{RP}^2} \\ = (N_0 \mp 2)\frac{\partial \tilde{\mathbf{R}}_{S^2}}{\partial S_0} + \sum_{i=1}^{K} N_i \left(\frac{\partial \tilde{\mathbf{R}}_{S^2}}{\partial \widetilde{S}_i} + \frac{\partial \tilde{\mathbf{R}}_{S^2}}{\partial \widetilde{S}_{-i}}\right) \pm \frac{2}{z}. \end{aligned} \tag{4.A.14}$$

Here $\tilde{\mathbf{R}}_{S^2}$ is the matrix model resolvent associated with the $U(\widetilde{N})$ theory. Using (4.A.12) and the relations $\mathbf{R}_{S^2} = R$, $\tilde{\mathbf{R}}_{S^2} = \widetilde{R}$ (Eq. (4.A.8)), we obtain

$$\mathbf{R}_{\mathbb{RP}^2}(z) = \mp\frac{1}{2}\frac{\partial}{\partial S_0}\mathbf{R}_{S^2}(z) \pm \frac{1}{2z}. \tag{4.A.15}$$

By expanding the resolvents around $z = \infty$ using (4.A.5) and comparing the coefficients, we obtain a relation between matrix model free energies:

$$j\frac{\partial \mathcal{F}_{\mathbb{RP}^2}}{\partial g_j} = \pm\frac{1}{2}\frac{\partial}{\partial S_0}\left(S\delta_{j0} + j\frac{\partial \mathcal{F}_{S^2}}{\partial g_j}\right) \mp \frac{1}{2}\delta_{j0}. \tag{4.A.16}$$

where $j = 0, 2, \cdots, 2(K+1)$. The $j = 0$ case is trivially satisfied since $S = S_0 + 2\sum_{i=1}^{K} S_i$ here, while the $j = 2, 4, \ldots, 2(K+1)$ cases lead to the first equation of (4.A.1), which we wanted to prove.

Next, consider $SO/Sp(N)$ theory with symmetric/antisymmetric tensor. The breaking pattern is $SO(N) \to \prod_{i=1}^{K} SO(N_i)$ or $Sp(N) \to \prod_{i=1}^{K} Sp(N_i)$ (Eq. (4.2.1)), where $N = \sum_{i=1}^{K} N_i$. It was shown in [24] that the resolvents of this SO/Sp theory is related to the resolvents $\tilde{R}(z)$ and $\widetilde{T}(z)$ of $U(\widetilde{N} \equiv N \mp 2K)$ theory with adjoint as follows:

$$R(z) = \tilde{R}(z), \qquad T(z) = \widetilde{T}(z) \pm \frac{d}{dz} \ln[W'(z)^2 + f_{K-1}(z)]. \tag{4.A.17}$$

The tree level superpotential of the $U(\widetilde{N})$ theory is related to the one for the $SO/Sp(N)$ theory as $W^U(z) = W^{SO/Sp}(z)$ (see (4.2.2) and (4.2.4)), and the breaking pattern is $U(\widetilde{N}) \to \prod_{i=1}^{K} U(N_i \mp 2)$. The glueball S_i of the $U(\widetilde{N})$ theory is taken to be the same as the glueball of the $SO/Sp(N)$ theory. In (4.A.17), $f_{K-1}(z)$ is a polynomial of degree $K-1$. Using (4.A.8), we can translate the relation (4.A.17) among gauge theory resolvents into a relation among matrix model resolvents:

$$\mathbf{R}_{\mathbb{RP}^2}(z) = \mp\frac{1}{2}\sum_{i=1}^{K} \frac{\partial}{\partial S_i}\mathbf{R}_{S^2}(z) \pm \frac{1}{4}\frac{d}{dz} \ln[W'(z)^2 + f_{K-1}(z)]. \tag{4.A.18}$$

In order to extract the relation between matrix model free energies, let us multiply (4.A.18) by z^j $(0 \le j \le K+1)$ and integrate over z along the contour C, introduced under (4.A.6), which encloses all the cuts around the critical points of $W(z)$. Taking

$$W'(z)^2 + f_{K-1}(z) = g_{K+1}^2 \prod_{i=1}^{K} (z - a_i^+)(z - a_i^-), \tag{4.A.19}$$

the branching points of the cuts are at $z = a_i^{\pm}$. The second term on the right hand side in (4.A.18) does not contribute to the contour integral unless $j = 0$:

$$\mp\frac{1}{4}\oint_C \frac{dz}{2\pi i}\sum_{i=1}^{K} z^j\left(\frac{1}{z-a_i^+}+\frac{1}{z-a_i^-}\right)$$
$$=\pm\frac{1}{4}\oint_C \frac{dw}{2\pi i}\sum_{i=1}^{K}\frac{1}{w^{j+1}}\left(\frac{1}{1-a_i^+ w}+\frac{1}{1-a_i^- w}\right)=\mp\frac{K}{2}\delta_{j0}, \quad (4.A.20)$$

where $w = 1/z$, because all the poles $w = 1/a_i^{\pm}$ are outside of the contour C (on the w-plane). On the contour C, we can use the Laurent expansion (4.A.5) to evaluate the contribution from the other terms, and the final result is

$$j\frac{\partial \mathcal{F}_{\mathbb{RP}^2}}{\partial g_j}=\pm\frac{1}{2}\sum_{i=1}^{K}\frac{\partial}{\partial S_i}\left(S\delta_{j0}+j\frac{\partial \mathcal{F}_{S^2}}{\partial g_j}\right)\mp\frac{K}{2}\delta_{j0}. \quad (4.A.21)$$

The $j = 0$ case is trivially satisfied, while the $1 \le j \le K+1$ cases lead to the second equation of (4.A.1), which we wanted to prove.

4.A.2 Computation of matrix model free energy: $SO/Sp(N)$ theory with adjoint

Let us consider $SO(\mathbf{N})$ matrix model which corresponds to $SO(N)$ gauge theory with adjoint. The tree level superpotential is taken to be quartic (4.5.1). The matrix variable $\mathbf{\Phi}$ in (4.A.2) is a real antisymmetric matrix and can be skew-diagonalized as

$$\mathbf{\Phi} \cong \mathrm{diag}[\lambda_1, \cdots, \lambda_{\mathbf{N}/2}] \otimes i\sigma^2. \quad (4.A.22)$$

By changing the integration variables from $\mathbf{\Phi}$ to λ_i, we obtain

$$\mathbf{Z} \sim \int \prod_{i=1}^{\mathbf{N}/2} d\lambda_i \prod_{i<j}^{\mathbf{N}/2}(\lambda_i^2-\lambda_j^2)^2\ e^{-\frac{1}{\mathbf{g}}\sum_{i=1}^{\mathbf{N}/2}\left(-\frac{m}{2}\lambda_i^2+\frac{g}{4}\lambda_i^4\right)}, \quad (4.A.23)$$

where $\prod_{i<j}^{\mathbf{N}/2}(\lambda_i^2-\lambda_j^2)^2$ is the Jacobian for this change of variables [78, 39]. The polynomial $-\frac{m}{2}\lambda^2+\frac{g}{4}\lambda^4$ has critical points at $\lambda = 0, \pm\sqrt{m/g}$, around which we

would like to do perturbative expansion. For this purpose, we separate λ's into two groups as

$$\lambda_i = \begin{cases} \lambda^{(0)}_{i_0} & i_0 = 1, \dots, \mathbf{N}_0/2, \\ \sqrt{m/g} + \lambda^{(1)}_{i_1} & i_1 = 1, \dots, \mathbf{N}_1, \end{cases} \tag{4.A.24}$$

with $\mathbf{N}_0 + 2\mathbf{N}_1 = \mathbf{N}$, corresponding to the classical supersymmetric vacuum with breaking pattern $SO(\mathbf{N}) \to SO(\mathbf{N}_0) \times U(\mathbf{N}_1)$. We would like to evaluate the matrix integral (4.A.23) perturbatively around $\lambda^{(0,1)} = 0$. If we expand the matrix model free energy in the coupling constant g as

$$\mathbf{F} = g f_1(\mathbf{N}_0, \mathbf{N}_1) + g^2 f_2(\mathbf{N}_0, \mathbf{N}_1) + \cdots , \tag{4.A.25}$$

the loop expansion tells us that $f_n(\mathbf{N}_0, \mathbf{N}_1)$ is a polynomial of degree $n+2$. Therefore, by performing the matrix integral by computer for small values of $\mathbf{N}_0$ and $\mathbf{N}_1$, one can determine the polynomial f_n. If we rewrite $\mathbf{N}_{0,1}$ in favor of $S_{0,1} = \mathbf{g}\mathbf{N}_{0,1}$, the expansion (4.A.25) arranges itself into the 't Hooft expansion (4.A.3), from which one can read off $\mathcal{F}_{S^2}$, $\mathcal{F}_{\mathbb{RP}^2}$, etc.

Following the procedure sketched above, we computed the matrix model free energy as

$$\begin{aligned} \mathcal{F}_{S^2} = {} & \left(\frac{1}{4}S_0^3 - 2S_0^2 S_1 + S_0 S_1^2\right)\alpha + \left(-\frac{9}{16}S_0^4 + 7S_0^3 S_1 - 9S_0^2 S_1^2 + 2S_0 S_1^3\right)\alpha^2 \\ & + \left(\frac{9}{4}S_0^5 - \frac{233}{6}S_0^4 S_1 + \frac{262}{3}S_0^3 S_1^2 - \frac{152}{3}S_0^2 S_1^3 + \frac{20}{3}S_0 S_1^4\right)\alpha^3 + \mathcal{O}(\alpha^4), \end{aligned} \tag{4.A.26}$$

where we defined $\alpha \equiv g/m^2$. We also checked explicitly that the relation (4.A.1) holds. Substituting (4.A.26) into the DV relation (4.5.3), we obtain the superpotential (4.5.5).

The $Sp(N)$ result is obtained similarly, with the result as in (4.A.1).

4.A.3 Computation of matrix model free energy: $Sp(N)$ theory with antisymmetric tensor

Consider the $Sp(\mathbf{N})$ matrix model which corresponds to $Sp(N)$ gauge theory with an antisymmetric tensor. The superpotential is taken to be quartic (4.5.12). The matrix variable $\mathbf{\Phi}$ satisfies $\mathbf{\Phi} = \mathbf{A}J$, $\mathbf{A}^T = -\mathbf{A}$. The "action" W_{tree} is given in (4.5.12). By a complexified $Sp(\mathbf{N})$ gauge rotation, the matrix $\mathbf{\Phi}$ can be brought to the form [21]

$$\Phi \cong \mathrm{diag}[\lambda_1, \ldots, \lambda_{\mathbf{N}/2}] \otimes \mathbb{1}_2, \qquad \lambda_i \in \mathbb{C}. \tag{4.A.27}$$

By changing the integration variables from Φ to λ_i, we obtain

$$Z \sim \int \prod_{i=1}^{\mathbf{N}/2} d\lambda_i \prod_{i<j}^{\mathbf{N}/2} (\lambda_i - \lambda_j)^4 \; e^{-\frac{1}{\mathbf{g}} \sum_{i=1}^{\mathbf{N}/2} \left(\frac{m}{2}\lambda_i^2 + \frac{g}{3}\lambda_i^3 \right)}. \tag{4.A.28}$$

where $\prod_{i<j}^{\mathbf{N}/2}(\lambda_i - \lambda_j)^4$ comes from the Jacobian for this change of variables. The polynomial $\frac{m}{2}\lambda^2 + \frac{g}{3}\lambda^3$ has two critical points $z = 0, -\frac{m}{g}$, around which we would like to do perturbative expansion. For this purpose, we separate λ's into two groups as

$$\lambda_i = \begin{cases} \lambda_{i_0}^{(1)} & i_1 = 1, \ldots, \mathbf{N}_1/2, \\ -m/g + \lambda_{i_1}^{(2)} & i_2 = 1, \ldots, \mathbf{N}_2/2, \end{cases} \tag{4.A.29}$$

with $\mathbf{N}_1 + \mathbf{N}_2 = \mathbf{N}$. This corresponds to the classical supersymmetric vacuum with breaking pattern $Sp(\mathbf{N}) \to Sp(\mathbf{N}_1) \times Sp(\mathbf{N}_2)$.

The matrix integral can be performed just the same way as for the $SO/Sp(N)$ theory with adjoint, as described in the last subsection. After substitution $S_{1,2} =$

$\mathbf{gN}_{1,2}$, we obtain

$$\begin{aligned}
\mathcal{F}_{S^2} = & \left(-\frac{S_1^3}{3} + \frac{S_2^3}{3} + \frac{5}{2}S_1^2 S_2 - \frac{5}{2}S_1 S_2^2\right)\alpha \\
& + \left(-\frac{4}{3}S_1^4 + \frac{91}{6}S_1^3 S_2 - \frac{59}{2}S_1^2 S_2^2 + \frac{91}{6}S_1 S_2^3 - \frac{4}{3}S_2^4\right)\alpha^2 \\
& + \left(-\frac{28}{3}S_1^5 + \frac{871}{6}S_1^4 S_2 - \frac{1318}{3}S_1^3 S_2^2 + \frac{1318}{3}S_1^2 S_2^3 - \frac{871}{6}S_1 S_2^4 + \frac{28}{3}S_2^5\right)\alpha^3 \\
& + \mathcal{O}(\alpha^4),
\end{aligned} \tag{4.A.30}$$

where $\alpha \equiv g^2/m^3$. We also checked explicitly that the relation (4.A.1) holds. Substituting (4.A.30) into (4.5.14), we obtain the glueball superpotential (4.5.16).

4.B Gauge theory calculation of superpotential

In this appendix, we compute the exact superpotential of the $\mathcal{N} = 1$ $SO/Sp(N)$ theory with adjoint in various vacua by considering factorization of the $\mathcal{N} = 2$ curve. This factorization method was developed in [10] for $U(N)$, and generalized in [79] to the case with unoccupied critical points (in other words, the $n < K$ case below). Inclusion of fundamentals was considered in [80, 81]. The generalization to SO/Sp gauge group, which discuss below, was given in [64, 71, 82, 83, 84, 25].

First consider $\mathcal{N} = 2$ $SO(N)$ theory broken to $\mathcal{N} = 1$ by the following polynomial tree level superpotential for the adjoint chiral superfield Φ:

$$W_{\text{tree}} = \frac{1}{2}\text{Tr}[W(\Phi)],$$
$$W(x) = \sum_{j=1}^{K+1} \frac{g_{2j}}{2j} x^{2j}, \qquad W'(x) = g_{2K+2}\, x \prod_{i=1}^{K}(x^2 - a_i^2). \tag{4.B.1}$$

The classical supersymmetric vacua are obtained by putting N_0 eigenvalues of Φ at $x = 0$ and N_i pairs of eigenvalues at $x = \pm a_i$, where $i = 1, \cdots, K$ and

$N_0 + 2\sum_{i=1}^{K} N_i = N$. In this vacuum the gauge group breaks as $SO(N) \to SO(N_0) \times \prod_{i=1}^{K} U(N_i)$. We allow some of N_i to vanish, i.e. we allow "unoccupied" critical points. Let the number of nonzero $N_{i\geq 1}$ be n. Then, the $\mathcal{N} = 2$ curve governing this $SO(N)$ theory factorizes as [64, 71, 82]:

$$y^2 = P_N^2(x) - 4x^4\Lambda^{2N-4} = \begin{cases} [xH_{N-2n-2}(x)]^2 F_{2(2n+1)}(x) & N_0 > 0, \\ H_{N-2n}(x)^2 F_{4n}(x) & N_0 = 0. \end{cases} \tag{4.B.2}$$

Here P, H and F are polynomials in x of the subscripted degree, which are invariant under $x \to -x$, i.e. they are actually polynomials in x^2. This factorization is required to have the appropriate number of independent, massless, monopoles and dyons, which must condense to eliminate some of the low-energy photons. The polynomial F is related to the tree level superpotential as

$$\begin{cases} F_{2(2n+1)}(x) = \frac{1}{g_{2K+2}^2} W'_{2K+1}(x)^2 + f_{2K}(x) & n = K, \\ F_{2(2n+1)}(x) Q_{2K-2n}(x)^2 = \frac{1}{g_{2K+2}^2} W'_{2K+1}(x)^2 + f_{2K}(x) & n < K \end{cases} \tag{4.B.3}$$

with some polynomial $Q_{2K-2n}(x)$, $f_{2K}(x)$ of the subscripted degrees, and $W'_{2K+1}(x)$ is as in (4.2.6). Equation (4.B.3) is for $N_0 > 0$, and for $N_0 = 0$ one must use the second equation with Q_{2K-2n} replaced by $Q_{2K-2n+2}$.

For $N_0 > 0$ we can write the solution of (4.B.2) in terms of that of the corresponding $U(N-2)$ breaking pattern, via:

$$P_N^{SO(N)}(x) = x^2 P_{N-2}^{U(N-2)}(x). \tag{4.B.4}$$

The low energy superpotential is given by

$$W_{\text{low}} = \frac{1}{2} \sum_{j=1}^{K+1} g_{2j} \langle u_{2j} \rangle, \tag{4.B.5}$$

where the $\langle u_{2j} \rangle$ are constrained to satisfy (4.B.2). Implementing this leads to the result that [83]:

$$\left\langle \mathrm{Tr}\frac{1}{x-\Phi} \right\rangle = \frac{d}{dx} \ln \left[P_N(x) + \sqrt{P_N(x)^2 - 4x^4\Lambda^{2N-4}} \right]. \tag{4.B.6}$$

Plugging back into (4.B.5) gives W_{low}. Note that the superpotential takes this simple form (4.B.5) only after one integrates out the monopoles and dyons, whose equation of motion led to the factorization constraint (4.B.2) [36, 37, 40].

The $Sp(N)$ theory can be solved similarly. The $\mathcal{N}=2$ curve factorizes in the vacuum with breaking pattern $Sp(N) \to Sp(N_0) \times \prod_{i=1}^{K} U(N_i)$ as [64, 71, 82]

$$\begin{aligned} y^2 = B_{N+2}(x)^2 - 4\Lambda^{2N+4} &= x^2 H_{N-2n}(x)^2 F_{2(2n+1)}(x), \\ B_{N+2}(x) &\equiv x^2 P_N(x) + 2\Lambda^{N+2}. \end{aligned} \tag{4.B.7}$$

The polynomial $F_{2(2n+1)}(x)$ is related to $W'(x)$ by (4.B.3). The mapping of the $Sp(N)$ theory to a $U(N+2)$ theory, as in (4.3.16), can be written as a solution of (4.B.7) in terms of solutions of the corresponding $U(N+2)$ factorization problem:

$$B_{N+2}(x) \equiv x^2 P_N^{Sp(N)}(x) + 2\Lambda^{N+2} = P_{N+2}^{U(N+2)}(x). \tag{4.B.8}$$

The Λ^{2N+2} shift in (4.B.8) is an $Sp(N)$ residual instanton effect, associated with the index of the embedding of the $U(N_i)$ factors in $Sp(N)$ [85, 75].

Again, the superpotential is given as in (4.B.5), subject to the constraint that $\langle u_j \rangle$ satisfy (4.B.7). Implementing this, the $\langle u_j \rangle$ can be obtained from $B_{N+2}(x)$ by [83]

$$\left\langle \mathrm{Tr}\frac{1}{x-\Phi} \right\rangle = \frac{d}{dx} \ln \left[B_{N+2}(x) + \sqrt{B_{N+2}(x)^2 - 4\Lambda^{2N+4}} \right]. \tag{4.B.9}$$

We now consider the exact W_{eff} for a few $SO/Sp(N)$ cases, to illustrate and clarify the general features[7]. We take quartic tree level superpotential

$$W_{\mathrm{tree}} = \frac{1}{2}\mathrm{Tr}W(\Phi), \qquad W(x) = \frac{m}{2}x^2 + \frac{g}{4}x^4, \tag{4.B.10}$$

[7]More SO/Sp examples can be found in [83].

which corresponds to $K = 1$.

4.B.1 $SO(N)$ unbroken

By the map (4.3.14), this maps to $U(N-2)$ unbroken, for which $P^{U(N-2)}(x) = 2\Lambda^{N-2}T_{N-2}(x/2\Lambda)$, with $T_N(x = \frac{1}{2}(t+t^{-1})) = \frac{1}{2}(t^N + t^{-N})$ a Chebyshev polynomial [86]. Thus, using (4.B.4), $P^{SO(N)}(x) = 2\Lambda^{N-2}x^2T_{N-2}(x/2\Lambda)$, as found in [40]. This then leads to [40]

$$\langle u_{2p}\rangle \equiv \frac{1}{2p}\langle \mathrm{Tr}\Phi^{2p}\rangle = \frac{N-2}{2p}\binom{2p}{p}\Lambda^{2p}. \tag{4.B.11}$$

In particular,

$$\langle u_2\rangle = (N-2)\Lambda^2, \quad \langle u_4\rangle = 3\left(\frac{N}{2}-1\right)\Lambda^4, \tag{4.B.12}$$

and the low-energy superpotential is $W_{\mathrm{low}} = \frac{1}{2}(m\langle u_2\rangle + g\langle u_4\rangle)$:

$$W_{\mathrm{low}} = \left(\frac{N}{2}-1\right)\left(m\Lambda^2 + \frac{3}{2}g\Lambda^4\right). \tag{4.B.13}$$

4.B.2 $SO(N) \to SO(2)\times U(N/2-1)$

By the map (4.3.14), this maps to $U(N-2) \to U(0)\times U(N/2-1)\times U(N/2-1)$. Using (4.3.14), the multiplication map of [10] for the $U(N-2)$ theory leads to a similar multiplication map for the $SO(N)$ theory, which was discussed in [83]. Using this, we can construct the solution to the factorization problem for general N in terms of that of say $N = 4$, i.e. $SO(4) \to SO(2)\times U(1)$. In this case, equation (4.B.2) is

$$y^2 = P_4^2 - 4x^4\Lambda^4 = x^2F_6. \tag{4.B.14}$$

The solution to this factorization problem is

$$P_4 = x^2(x^2-a^2), \quad F_6 = x^2[(x^2-a^2)^2 - 4\Lambda^2], \tag{4.B.15}$$

from which we can see the breaking pattern $SO(4) \rightarrow SO(2) \times U(1)$. Using (4.B.6) gives

$$u_2 = a^2, \quad u_4 = \frac{a^4}{2} + \Lambda^4. \tag{4.B.16}$$

Further, the condition (4.B.3)

$$F_6 = \frac{1}{g^2} W_3'^2 + f_2 \tag{4.B.17}$$

leads to

$$a^2 = -\frac{m}{g}, \quad f_2 = -4\Lambda^2 x^2. \tag{4.B.18}$$

The solution for general $SO(N) \rightarrow SO(2) \times U(N/2-1)$, the multiplication map gives the solution to the factorization problem as $P_N(x) = 2x^2\Lambda^{N-2}T_{N/2-1}((x^2 - a^2)/2\Lambda^2)$, with $T_{N/2-1}$ the Chebyshev polynomial defined above. The effect is to rescale u_2, u_4, and hence W_{low} by an overall factor of $N/2-1$:

$$W_{\text{low}} = \left(\frac{N}{2} - 1\right)\left(-\frac{m^2}{4g} + \frac{1}{2}g\Lambda^4\right). \tag{4.B.19}$$

This agrees with the result (4.5.10).

4.B.3 $SO(4) \rightarrow U(2)$

More generally, we could consider the breaking pattern $SO(N) \rightarrow U(N/2)$. The map of (4.3.14) is less useful here, when $N_0 = 0$, since it suggests mapping to $U(N-2) \rightarrow U(-2) \times U(N/2) \times U(N/2)$ and the $U(-2)$ needs to be interpreted. In general, this breaking pattern leads to a complicated $W_{\text{low}}(\Lambda)$. We will here illustrate the case $SO(4) \rightarrow U(2)$, corresponding to $N = 4,\ N_0 = 0,\ n = 1,\ K = 1$. Equation (4.B.2) is

$$y^2 = P_4^2 - 4x^4\Lambda^4 = H_2^2 F_4. \tag{4.B.20}$$

The solution to this factorization problem is

$$P_4 = (x^2 - a^2)^2 + 2\Lambda^2 x^2, \quad H_2 = x^2 - a^2, \quad F_4 = (x^2 - a^2)^2 + 4\Lambda^2 x^2. \quad (4.B.21)$$

In the classical $\Lambda \to 0$ limit, this shows $P_4 \to (x^2 - a^2)^2$, implying the breaking pattern $SO(4) \to U(2)$. (4.B.6) gives

$$u_2 = 2(a^2 - \Lambda^2), \quad u_4 = (a^4 - 2\Lambda^2)^2 - \Lambda^4. \quad (4.B.22)$$

Further, the condition (4.B.3)

$$F_4 x^2 = \frac{1}{g^2} W_3'^2 + f_2 \quad (4.B.23)$$

leads to

$$a^2 = -\frac{m}{g} + 2\Lambda^2, \quad f_2 = 4\Lambda^2 x^2 \Big(-\frac{m}{g} + \Lambda^2 \Big). \quad (4.B.24)$$

Therefore the exact superpotential is

$$W_{\mathrm{low}} = -\frac{m^2}{2g} + m\Lambda^2 + \frac{1}{2} g\Lambda^4. \quad (4.B.25)$$

4.B.4 $Sp(4) \to U(2)$, $Sp(2) \times U(1)$

This corresponds to $N = 4,\ n = 1,\ K = 1$. Equations (4.B.7) and (4.B.3) are

$$y^2 = B_6^2 - 4\Lambda^{12} = x^2 H_2^2 F_6, \qquad F_6 = \frac{1}{g^2} W_3'^2 + f_2. \quad (4.B.26)$$

This factorization problem is solved by [83]:

$$P_4 = (x^2 - a^2)^2 + \frac{4\Lambda^6}{a^4}(x^2 - 2a^2), \qquad \frac{m}{g} = -a^2 + \frac{4\Lambda^6}{a^4} \quad (4.B.27)$$

From (4.B.9), we obtain

$$u_2 = 2a^2 - \frac{4\Lambda^6}{a^4}, \qquad u_4 = a^4 + \frac{8\Lambda^{12}}{a^8}. \quad (4.B.28)$$

This solution continuously connects two classically different vacua with breaking pattern $Sp(4) \to U(2)$ and $Sp(4) \to Sp(2) \times U(1)$. Correspondingly there are two ways to take the classical limit: i) $\Lambda \to 0$ with a fixed, or ii) $\Lambda, a \to 0$ with $w = 2\Lambda^3/a^2$ fixed. In these limits, $P_4(x)$ goes to i) $(x^2 - a^2)^2$ or ii) $x^2(x^2 + w^2)$, showing the aforementioned breaking pattern.

In the $Sp(4) \to U(2)$ case, we solve the second equation of (4.B.27) with the condition $a^2 \to -m/g$ as $\Lambda \to 0$. The solution is

$$a^2 = -\frac{m}{g} + \frac{4g^2\Lambda^6}{m^2} + \frac{32g^5\Lambda^{12}}{m^5} + \frac{448g^8\Lambda^{18}}{m^8} + \cdots . \tag{4.B.29}$$

From (4.B.28) and (4.B.5), one obtains the exact superpotential:

$$W_{\mathrm{low}} = -\frac{m^2}{2g} - \frac{2g^2\Lambda^6}{m} - \frac{4g^5\Lambda^{12}}{m^4} - \frac{32g^8\Lambda^{18}}{m^7} + \cdots . \tag{4.B.30}$$

In the $Sp(4) \to Sp(2) \times U(1)$ case, we solve the second equation of 4.B.27 with the condition $w^2 \to m/g$ as $\Lambda \to 0$. It is

$$w = \frac{m^{1/2}}{g^{1/2}} + \frac{g\Lambda^3}{m} - \frac{3g^{5/2}\Lambda^6}{2m^{5/2}} + \frac{4g^4\Lambda^9}{m^4} + \cdots . \tag{4.B.31}$$

From (4.B.28) and (4.B.5), one obtains the exact superpotential:

$$W_{\mathrm{low}} = -\frac{m^2}{4g} + 2m^{1/2}g^{1/2}\Lambda^3 + \frac{g^2\Lambda^6}{m} - \frac{g^{7/2}\Lambda^9}{m^{5/2}} + \cdots . \tag{4.B.32}$$

CHAPTER 5

Adding flavors

We present two results concerning the relation between poles and cuts by using the example of $\mathcal{N} = 1$ $U(N_c)$ gauge theories with matter fields in the adjoint, fundamental and anti-fundamental representations. The first result is the on-shell possibility of poles, which are associated with flavors and on the second sheet of the Riemann surface, passing through the branch cut and getting to the first sheet. The second result is the generalization of hep-th/0311181 (Intriligator, Kraus, Ryzhov, Shigemori, and Vafa) to include flavors. We clarify when there are closed cuts and how to reproduce the results of the strong coupling analysis by matrix model, by setting the glueball field to zero from the beginning. We also make remarks on the possible stringy explanations of the results and on generalization to $SO(N_c)$ and $USp(2N_c)$ gauge groups.

5.1 Introduction

String theory can be a powerful tool to understand four dimensional supersymmetric gauge theory which exhibits rich dynamics and allows an exact analysis. In [35], using the generalized Konishi anomaly and matrix model [5, 6, 7], $\mathcal{N} = 1$ supersymmetric $U(N_c)$ gauge theory with matter fields in the adjoint, fundamental and anti-fundamental representations was studied. The resolvents in the quantum theory live on the two-sheeted Riemann surface defined by the matrix

model curve. Their quantum behavior is characterized by the structure around the branch cuts and poles, which are related to the RR flux contributions in the Calabi–Yau geometry and flavor fields, respectively. A pole associated with flavor on the first sheet is related to the Higgs vacua (corresponding to classical nonzero vacuum expectation value of the fundamental) while a pole on the second sheet is related to the pseudo-confining vacua where the classically vanishing vacuum expectation value of the fundamental gets nonzero values due to quantum correction.

It is known [35] that Higgs vacua and pseudo-confining vacua, which are distinct in the classical theory, are smoothly transformed into each other in the quantum theory. This transition is realized on the Riemann surface by moving poles located on the second sheet to pass the branch cuts and enter the first sheet. This process was analyzed in [35] at the off-shell level by fixing the value of glueball fields during the whole process. However, in an on-shell process, the position of poles and the width and position of branch cuts are correlated (when the flavor poles are moved, the glueball field is also changed). It was conjectured in [35] that for a given branch cut, there is an upper bound for the number of poles (the number of flavors) which can pass through the cut from the second sheet to the first sheet.

Our first aim of this paper is to confirm this conjecture and give the corresponding upper bound for various gauge groups (in particular, we will concentrate on the $U(N_c)$ gauge group). The main result is that if $N_f \geq N_c$, the poles will not be able to pass through the cut to the first sheet where N_c is the effective fluxes associated with the cut (and can be generalized to other gauge groups).

Another important development was made in [87], which was inspired by [23]. In [87], which we will refer to as IKRSV, it was shown that, to correctly compute

the prediction of string theory (matrix model), it is crucial to determine whether the glueball is really a good variable or not. A prescription was given, regarding when a glueball field corresponding to a given branch cut should be set to zero before extremizing the off-shell glueball superpotential. The discussion of IKRSV was restricted to $\mathcal{N} = 1$ gauge theories with an adjoint and no flavors, so the generalization to the case with fundamental flavors is obviously the next task.

Our second aim of this paper is to carry out this task. The main result is the following. Assuming N_f poles around a cut associated with gauge group $U(N_{c,i})$, when $N_f \geq N_{c,i}$ there are situations in which we should set $S_i = 0$ in matrix model computations. More concretely, situations with $S_i = 0$ belong to either of the following two branches: the baryonic branch for $N_{c,i} \leq N_f < 2N_{c,i}$, or the $r = N_{c,i}$ non-baryonic branch for $N_f \geq 2N_{c,i}$. Moreover, when $S_i = 0$, the gauge group is completely broken and there should exist some extra, charged massless field which is not incorporated in matrix model.

In section 5.2, as background, we review basic materials for $\mathcal{N} = 1$ supersymmetric $U(N_c)$ gauge theory with an adjoint chiral superfield, and N_f flavors of quarks and anti-quarks. The chiral operators and the exact effective glueball superpotential are given. We study the vacuum structure of the gauge theory at classical and quantum levels. We review also the main results of IKRSV. In addition to all these reviews, we present our main motivations of this paper.

In section 5.3, we apply the formula for the off-shell superpotential obtained in [35] to the case with quadratic tree level superpotential, and solve the equation of motion derived from it. We consider what happens if one moves N_f poles associated with flavors on the second sheet through the cut onto the first sheet, on-shell. Also, in subsection 5.3.4, we briefly touch the matter of generalizing IKRSV in the one cut model.

In section 5.4, we consider cubic tree level superpotential. On the gauge theory side, the factorization of the Seiberg–Witten curve provides an exact superpotential. We reproduce this superpotential by matrix model, by extremizing the effective glueball superpotential with respect to glueball fields after setting the glueball field to zero when necessary. We present explicit results for $U(3)$ theory with all possible breaking patterns and different number of flavors ($N_f = 1, 2, 3, 4$, and 5).

In section 5.5, after giving concluding remarks, we repeat the procedure we did in previous sections for $SO(N_c)/USp(2N_c)$ theories, briefly.

In the appendix, we present some proofs and detailed calculations which are necessary for the analysis in section 5.4.

Since string theory results in the dual Calabi–Yau geometry are equivalent to the matrix model results, we refer to them synonymously through the paper. There exist many related works to the present paper. For a list of references, we refer the reader to [19].

5.2 Background

In this section, we will summarize the relevant background needed for the study of $\mathcal{N} = 1$ supersymmetric gauge theory with matter fields.

5.2.1 The general picture of matrix model with flavors

The generalized Konishi anomaly interpretation to the matrix model approach for $\mathcal{N} = 1$ supersymmetric gauge theory with flavors was given in [47, 35]. Here we make only a brief summary on some points we will need.

Let us consider $\mathcal{N} = 1$ supersymmetric $U(N_c)$ gauge theory, coupled to an

adjoint chiral superfield Φ, N_f fundamentals Q^f, and N_f anti-fundamentals $\widetilde{Q}_{\tilde{f}}$. The tree level superpotential is taken to be

$$W_{\rm tree} = \operatorname{Tr} W(\Phi) + \sum_{f,\tilde{f}} \widetilde{Q}_{\tilde{f}}\, m^{\tilde{f}}{}_f(\Phi) Q^f \ , \tag{5.2.1}$$

where the function $W(z)$ and the matrix $m^{\tilde{f}}_f(z)$ are polynomials

$$W(z) = \sum_{k=0}^{n} \frac{g_k z^{k+1}}{k+1}, \qquad m^{\tilde{f}}{}_f(z) = \sum_{k=1}^{l+1} (m_k)^{\tilde{f}}{}_f z^{k-1} \ .$$

Classically we can have the "pseudo-confining vacua" where the vacuum expectation values of Q, $\widetilde{Q}$ are zero, or the "Higgs vacua" where the vacuum expectation values of Q, $\widetilde{Q}$ are nonzero so that the total rank of the remaining gauge groups is reduced. These two vacua, which seem to have a big difference classically, are not fundamentally distinguishable from each other in the quantum theory and in fact can be continuously transformed into each other, as we will review shortly, in the presence of flavors [35].

Supersymmetric vacua of gauge theory are characterized by the vacuum expectation values of chiral operators [88]. They are nicely packaged into the following functions called resolvents [35, 47]: [1]

$$T(z) = \left\langle \operatorname{Tr} \frac{1}{z-\Phi} \right\rangle, \tag{5.2.2}$$

$$R(z) = -\frac{1}{32\pi^2} \left\langle \operatorname{Tr} \frac{W_\alpha W^\alpha}{z-\Phi} \right\rangle, \tag{5.2.3}$$

$$M(z)^f{}_{\tilde{f}} = \left\langle \widetilde{Q}_{\tilde{f}} \frac{1}{z-\Phi} Q^f \right\rangle \tag{5.2.4}$$

where W_α is (the lowest component of) the field strength superfield. Classically, $R(z)$ vanishes while $T(z)$, $M(z)$ have simple poles on the complex z-plane at

[1] We set $w_\alpha(z) \equiv \frac{1}{4\pi} \left\langle \operatorname{Tr} \frac{W_\alpha}{z-\Phi} \right\rangle$ to zero because in supersymmetric vacua $w_\alpha(z) = 0$.

infinity and at the eigenvalues of Φ. Each eigenvalue of Φ is equal to one of zeros of $W'(z)$ or $B(z)$, where

$$W'(z) = g_n \prod_{i=1}^{n} (z - a_i), \qquad B(z) \equiv \det m(z) = B_L \prod_{I=1}^{L} (z - z_I). \tag{5.2.5}$$

In the pseudo-confining vacuum, every eigenvalue of Φ is equal to a_i for some i. On the other hand, in the Higgs vacuum, some eigenvalues of Φ are equal to z_I for some I.

In the quantum theory, the resolvents (5.2.2), (5.2.3) and (5.2.4) are determined by the generalized Konishi anomaly equations [79, 47, 35]:

$$\begin{aligned}
[W'(z)T(z)]_- + \mathrm{Tr}[m'(z)M(z)]_- &= 2R(z)T(z), \\
[W'(z)R(z)]_- &= R(z)^2, \\
[(M(z)m(z))^{f'}{}_f]_- &= R(z)\delta_f^{f'}, \\
[(m(z)M(z))^{\tilde{f}'}{}_{\tilde{f}}]_- &= R(z)\delta_{\tilde{f}}^{\tilde{f}'},
\end{aligned} \tag{5.2.6}$$

where the notation $[\ ]_-$ means to drop the nonnegative powers in a Laurent expansion in z. From the second equation of (5.2.6), one obtains [17]

$$R(z) = \frac{1}{2}\left(W'(z) - \sqrt{W'(z)^2 + f(z)}\right),$$

where $f(z)$ is a polynomial of degree $(n-1)$ in z. This implies that in the quantum theory the zeros $z = a_i$ $(i = 1, 2, \ldots, n)$ of $W'(z)$ are blown up into cuts A_i along intervals[2] $[a_i^-, a_i^+]$ by the quantum effect represented by $f(z)$, and the resolvents (5.2.2)–(5.2.4) are defined on a double cover of the complex z-plane branched at the roots $a_i^\pm$ of $W'(z)^2 + f(z)$. This double cover of the z-plane can be thought of as a Riemann surface Σ described by the matrix model curve

$$\Sigma: \quad y_m^2 = W'(z)^2 + f(z). \tag{5.2.7}$$

[2] $a_i^\pm$ are generally complex and in such cases we take A_i to be a straight line connecting a_i^- and a_i^+. Note that there is no physical meaning to the choice of the cut; it can be any path connecting a_i^- and a_i^+.

This curve is closely related to the factorization form of $\mathcal{N} = 2$ curve in the strong coupling analysis.

Every point z on the z-plane is lifted to two points on the Riemann surface Σ which we denote by q and $\widetilde{q}$ respectively. For example, z_I is lifted to q_I on the first sheet and $\widetilde{q}_I$ on the second sheet. We write the projection from Σ to the z-plane as $z_I = z(q_I) = z(\widetilde{q}_I)$, following the notation of [35].

The classical singularities of the resolvents $T(z)$, $M(z)$ are modified in the quantum theory to the singularities on Σ, as follows. For $T(z)$, the classical poles at z_I are lifted to poles at q_I or $\widetilde{q}_I$, depending on which vacuum the theory is in, while the classical poles at a_i with residue $N_{c,i}$ are replaced by cuts with periods $\frac{1}{2\pi i}\oint_{A_i} T(z)dz = N_{c,i}$. For $M(z)$, the classical poles at z_I are also lifted to poles at q_I or $\widetilde{q}_I$. More specifically, by solving the last two equations of (5.2.6), one can show [35]

$$M(z) = R(z)\frac{1}{m(z)} - \sum_{I=1}^{L} \frac{(1-r_I)R(q_I)}{(z-z_I)} \frac{1}{2\pi i} \oint_{q_I} \frac{dx}{m(x)} - \sum_{I=1}^{L} \frac{r_I R(\widetilde{q}_I)}{(z-z_I)} \frac{1}{2\pi i} \oint_{\widetilde{q}_I} \frac{dx}{m(x)}, \tag{5.2.8}$$

where $(q_I, \widetilde{q}_I)$ are the lift of z_I to the first sheet and to the second sheet of Σ, and $r_I = 0$ for poles on the second sheet and $r_I = 1$ for poles on the first sheet. Furthermore, for $T(z)$, by solving the first equation of (5.2.6),

$$T(z) = \frac{B'(z)}{2B(z)} - \sum_{I=1}^{L} \frac{(1-2r_I)y(q_I)}{2y(z)(z-z_I)} + \frac{c(z)}{y(z)}, \tag{5.2.9}$$

where

$$c(z) = \left\langle \mathrm{Tr}\frac{W'(z) - W'(\Phi)}{z-\Phi} \right\rangle - \frac{1}{2}\sum_{I=1}^{L} \frac{W'(z) - W'(z_I)}{z-z_I}. \tag{5.2.10}$$

Practically it is hard to use (5.2.10) to obtain $c(z)$ and we use the following

condition instead:

$$\frac{1}{2\pi i}\oint_{A_i} T(z)dz = N_{c,i}. \tag{5.2.11}$$

Finally, the exact, effective glueball superpotential is given by [35]

$$\begin{aligned}
W_{\text{eff}} = & -\frac{1}{2}\sum_{i=1}^{n} N_{c,i}\int_{\widehat{B}_i^r} y(z)dz \\
& -\frac{1}{2}\sum_{I=1}^{L}(1-r_I)\int_{\widetilde{q}_I}^{\widetilde{\Lambda}_0} y(z)dz - \frac{1}{2}\sum_{I=1}^{L} r_I\int_{q_I}^{\widetilde{\Lambda}_0} y(z)dz \\
& +\frac{1}{2}(2N_c - L)W(\Lambda_0) + \frac{1}{2}\sum_{I=1}^{L} W(z_I) \\
& -\pi i(2N_c - L)S + 2\pi i\tau_0 S + 2\pi i\sum_{i=1}^{n-1} b_i S_i,
\end{aligned} \tag{5.2.12}$$

where

$$2\pi i\tau_0 = \log\left(\frac{B_L\Lambda^{2N_c-N_f}}{\Lambda_0^{2N_c-L}}\right). \tag{5.2.13}$$

Here, Λ_0 is the cut-off of the contour integrals, Λ is the dynamical scale, $S \equiv \sum_{i=1}^{n} S_i$, and $b_i \in \mathbb{Z}$. $\widehat{B}_i^r$ is the regularized contour from $\widetilde{\Lambda}_0$ to Λ_0 through the i-th cut and Λ_0 and $\widetilde{\Lambda}_0$ are the points on the first sheet and on the second sheet, respectively. The glueball field is defined as

$$S_i = \frac{1}{2\pi i}\oint_{A_i} R(z)dz.$$

In the above general solutions (5.2.8), (5.2.9), we have $r_I = 1$ or $r_I = 0$, depending on whether the pole is on the first sheet or on the second sheet of Σ, respectively. The relation between these choices of r_I and the phase of the system is as follows. Let us start with all $r_I = 0$, i.e., all the poles are on the second sheet. This choice corresponds to the pseudo-confining vacua where the gauge group is broken as $U(N_c) \to \prod_{i=1}^{n} U(N_{c,i})$ with $\sum_{i=1}^{n} N_{c,i} = N_c$. Now let us move

a single pole through, for example, the n-th cut to the first sheet. This will break the gauge group as $\prod_{i=1}^{n} U(N_{c,i}) \to \prod_{i=1}^{n-1} U(N_{c,i}) \times U(N_{c,n}-1)$. Note that the rank of the last factor is now $(N_{c,n}-1)$ so that $\sum_{i=1}^{n} N_{c,i} = N_c - 1 < N_c$. Namely, the gauge group is Higgsed down. In this way, by passing poles through cuts, one can go continuously from the pseudo-confining phase to the Higgs phase, as advocated before.

However, if we consider this process of passing poles through a cut to the first sheet *on-shell*, then there should be an obstacle at a certain point. For example, if initially we have $N_{c,n} = 1$, after passing a pole we would end up with an $U(0)$. This sudden jump of the number of $U(1)$'s in the low energy gauge theory is not a smooth physical process, because the number of massless particles (photons) changes discontinuously. So we expect some modifications to the above picture. In [35], it was suggested that in an on-shell process, the n-th cut will close up in such a situation so that the pole cannot pass through. It is one of our motivations to show that this is indeed true. More precisely, the cut does not close up completely and the pole can go through a little bit further to the first sheet and then will be bounced back to the second sheet.

5.2.2 The vacuum structure

In the last subsection we saw that different distributions of poles over the first and the second sheets correspond to different phases of the theory. In this subsection we will try to understand this vacuum structure of the gauge theory at both classical and quantum levels for a specific model (for more details, see [89, 90, 91, 81, 84, 25]). For simplicity we will focus on $U(N_c)$ theory with N_f flavors and

the following tree level superpotential [3]

$$W_{\text{tree}} = \frac{1}{2} m_A \operatorname{Tr} \Phi^2 - \sum_{I=1}^{N_f} \widetilde{Q}_I (\Phi + m_f) Q^I. \tag{5.2.14}$$

This corresponds to taking polynomials in (5.2.1) as

$$W(z) = \frac{m_A}{2} z^2, \qquad m^{\tilde{I}}{}_I(z) = -(z + m_f)\delta_I^{\tilde{I}}.$$

All N_f flavors have the same mass m_f, and the mass function defined in (5.2.5) is given by

$$B(z) = (-1)^{N_f} (z + m_f)^{N_f}.$$

Therefore, poles associated with flavors are located at

$$z_I = -m_f \equiv z_f, \qquad I = 1, 2, \ldots, N_f \ . \tag{5.2.15}$$

In the quantum theory, some of these poles are lifted to q_f on the first sheet and others are lifted to $\widetilde{q}_f$ on the second sheet.

The D- and F-flatness for the superpotential (5.2.14) is given by

$$\begin{aligned}
&0 = [\Phi, \Phi^\dagger], \quad 0 = QQ^\dagger - \widetilde{Q}^\dagger \widetilde{Q}, \\
&0 = m_A \Phi - Q\widetilde{Q}, \quad 0 = (\Phi + m_f) Q = \widetilde{Q}(\Phi + m_f).
\end{aligned}$$

Solutions are a little different for $m_f \neq 0$ and $m_f = 0$, because $m_f = 0$ is the root of $W'(z) = z$. The case of $W'(-m_f) = 0$ was discussed in [89, 81] which we will refer to as the classically massless case.

[3] We used the convention of [92] for the normalization of the second term. Different choices are related to each other by redefinition of $\widetilde{Q}$ and Q.

In the $m_f \neq 0$ case, the solution is given by

$$\Phi = \begin{bmatrix} -m_f I_{K\times K} & 0 \\ 0 & 0_{(N_c-K)\times(N_c-K)} \end{bmatrix},$$

$$Q = \begin{bmatrix} A_{K\times K} & 0 \\ 0 & 0_{(N_c-K)\times(N_f-K)} \end{bmatrix}, \quad {}^t\widetilde{Q} = \begin{bmatrix} {}^t\widetilde{A}_{K\times K} & \widetilde{B}_{K\times(N_f-K)} \\ 0 & 0_{(N_c-K)\times(N_f-K)} \end{bmatrix} \tag{5.2.16}$$

with

$$-m_f m_A I_{K\times K} = A\,{}^t\widetilde{A}, \qquad AA^\dagger = \widetilde{A}^\dagger\widetilde{A} + \widetilde{B}^*\,{}^t\widetilde{B}.$$

The gauge group is Higgsed down to $U(N_c - K)$ where

$$K_{m_f\neq 0} \leq \min\left(N_c, N_f\right).$$

To understand the range of $K_{m_f\neq 0}$, first note that the Φ breaks the gauge group as $U(N_c) \to U(K)\times U(N_c-K)$. Now the $U(K)$ factor has effectively N_f massless flavors and because $\langle\widetilde{Q}Q\rangle \neq 0$, $U(K)$ is further Higgsed down to $U(0)$.

For $m_f = 0$, we have $\Phi = 0$ and $Q, \widetilde{Q}$ are still of the above form (5.2.16) with one special requirement: $\widetilde{A} = 0$. Because of this we have

$$K_{m_f=0} \leq \min\left(N_c, \left[\frac{N_f}{2}\right]\right),$$

where [ ] means the integer part. The integer $K_{m_f=0}$ precisely corresponds to the r-th branch discussed in [89, 81]. The $m_f = 0$ case is different from the $m_f \neq 0$ case as follows. First, Φ does not break the $U(N_c)$ gauge group, i.e., $U(N_c) \to U(N_c)$. Secondly, the r-th branch is the intersection of the Coulomb branch in which $\langle\widetilde{Q}Q\rangle = 0$ and the Higgs branch in which $\langle\widetilde{Q}Q\rangle \neq 0$, whereas for $m_f \neq 0$ the vacuum expectation value $\langle\widetilde{Q}Q\rangle$ must be nonzero and the gauge group must be Higgsed down. For these reasons, $K_{m_f\neq 0}$ and $K_{m_f=0}$ have different ranges.

The above classical classification of r-th branches is also valid in the quantum theory (including the baryonic branch).

The quantum r-th branch can also be discussed by using the Seiberg–Witten curve. In the r-th branch, the curve factorizes as

$$\begin{aligned} y_{\mathcal{N}=2}^2 &= P_{N_c}(x)^2 - 4\Lambda^{2N_c-N_f}(x+m_f)^{N_f} \\ &= (x+m_f)^{2r}\left[P_{N_c-r}^2(x) - 4\Lambda^{2N_c-N_f}(x+m_f)^{N_f-2r}\right]. \end{aligned}$$

Because $N_c - r \geq 0$ (coming from $P_{N_c-r}(x)$) and $N_f - 2r \geq 0$ (coming from the last term), we have $r \leq N_c$ and $r \leq N_f/2$, which leads to the range

$$r \leq \min\left(N_c, \left[\frac{N_f}{2}\right]\right). \tag{5.2.17}$$

The relation between this classification of the Seiberg–Witten curve and the above classification of r-branches, in the $m_f \neq 0$ and $m_f = 0$ cases, is as follows. In the $m_f = 0$ case, we have one-to-one correspondence where the r is identified with $K_{m_f=0}$. In the $m_f \neq 0$ case, for a given r of the curve, there exist two cases: either $K_{m_f\neq 0} = r$ for $K_{m_f\neq 0} \leq [N_f/2]$, or $K_{m_f\neq 0} = N_f - r$ for $K_{m_f\neq 0} \geq [N_f/2]$.

5.2.3 The work of IKRSV

Now we discuss another aspect of the model. In the above, we saw that there is a period condition (5.2.11) for $T(z)$. So, if for the i-th cut we have $\oint_{A_i} T(z)dz = N_{c,i} = 0$, then it seems that, in the string theory realization of the gauge theory, there is no RR flux provided by D5-branes through this cut and the cut is closed. Because of this, it seems that we should set the corresponding glueball field $S_i = 0$. Based on this naive expectation, Ref. [23] calculated the effective superpotential of $USp(2N_c)$ theory with an antisymmetric tensor by the matrix model, which turned out to be different from the known results obtained by holomorphy and symmetry arguments (later Refs. [31, 30] confirmed this discrepancy).

This puzzle intrigued several papers [15, 24, 87, 25, 26, 29, 27, 28]. In particular, in [24], it was found that although $N_{c,i} = 0$, we cannot set $S_i = 0$. The reason became clear by later studies. Whether a cut closes or not is related to the total RR flux which comes from both D5-branes and orientifolds. For $USp(2N_c)$ theory with antisymmetric tensor, although the RR flux from D5-branes is zero, there exists RR flux coming from the orientifold with positive RR charges, thus the cut does not close. That the cut does not close can also be observed from the Seiberg–Witten curve [25] where for such a cut, we have two single roots in the curve, instead of a double root. All these results were integrated in [87] for $\mathcal{N} = 1$ gauge theory with adjoint. Let us define $\widehat{N}_c = N_c$ for $U(N_c)$, $N_c/2 - 1$ for $SO(N_c)$ and $2N_c + 2$ for $USp(2N_c)$. Then the conclusion of [87] can be stated as

> If $\widehat{N}_c > 0$, we should include S_i and extremize $W_{\text{eff}}(S_i)$ with respect to it. On the other hand, if $\widehat{N}_c \leq 0$, we just set $S_i = 0$ instead.

In [87], it was argued that this prescription of setting $S_i = 0$ can be explained in string theory realization by considering an extra degree of freedom which corresponds to the D3-brane wrapping the blown up S^3 and becomes massless in the $S \to 0$ limit [60]. Our second motivation of this paper is to generalize this conclusion to the case with flavors. We will discuss the precise condition when one should set $S_i = 0$ in order to get agreement with the gauge theory result, in the case with flavors.

5.2.4 Prospects from the strong coupling analysis

Before delving into detailed calculations, let us try to get some general pictures from the viewpoint of factorization of the Seiberg–Witten curve. Since we hope to generalize IKRSV, we are interested in the case where some S_i vanish. Because

S_i is related to the size of a cut in the matrix model curve, which is essentially the same as the Seiberg–Witten curve, we want some cuts to be closed in the Seiberg–Witten curve. Namely, we want a double root in the factorization of the curve, instead of two single roots.

For $U(N_c)$ theory with N_f flavors of the same mass $m_f = -z_f$, tree level superpotential (5.2.1), and breaking pattern $U(N_c) \to \prod_{i=1}^{n} U(N_{c,i})$, the factorization form of the curve is [81]

$$\begin{aligned} P_{N_c}(z)^2 - 4\Lambda^{2N_c - N_f}(z - z_f)^{N_f} &= H_{N-n}(z)^2 F_{2n}(z), \\ F_{2n}(z) &= W'(z)^2 + f_{n-1}(z), \end{aligned} \tag{5.2.18}$$

where the degree $2n$ polynomial $F_{2n}(z) = W'(z)^2 + f_{n-1}(z)$ generically has $2n$ single roots. How can we have a double root instead of two single roots?

For a given fixed mass, for example $z_f = a_1$, there are three cases where we have a double root, as follows. (a) There is no $U(N_{c,1})$ group factor associated with the root a_1, namely $N_{c,1} = 0$. (b) The $U(N_{c,1})$ factor is in the baryonic branch. This can happen for $N_{c,1} \leq N_f < 2N_{c,1}$. (c) The $U(N_{c,1})$ factor is in the r-th non-baryonic branch with $r = N_{c,1}$. This can happen only for $N_f \geq 2N_{c,1}$. Among these three cases, (b) and (c) [81] are new for theories with flavors, and will be the focus of this paper. However, it is worth pointing out that the factorization form in the cases (b) and (c) are not the one given in (5.2.18) for a fixed mass, but the one given in (5.A.3).

Can we keep the factorization form (5.2.18) while having an extra double root? We can, but instead of a fixed mass we must let the mass "floating," which means the following. There will be multiple solutions to the factorization form (5.2.18), and for any given solution the $2n$ single roots of $F_{2n}(x)$, denoted by $a_i^{\pm}$, $i = 1, \ldots, n$, are functions of z_f. Now, we tune z_f so that $a_i^+(z_f) = a_i^-(z_f)$, i.e., so that two single roots combine into one double root. Since for different

solutions this procedure will lead to different values of z_f, we call this situation the "floating" mass.

Now we have two ways to obtain extra double roots: one is with a fixed mass, but to go to the baryonic or the $r = N_{c,1}$ non-baryonic branch, while the other is to start with a general non-baryonic branch but using a floating mass. In fact it can be shown that these two methods are equivalent to each other when the double root is produced. In the calculations in section 5.4 we will use the floating mass to check our proposal.

5.3 One cut model—quadratic tree level superpotential

In this section we will study whether a cut closes up if one tries to pass too many poles through it. If the poles are near the cut, the precise form of the tree level superpotential (namely the polynomials $W(z)$, $m^{\tilde{f}}{}_f(z)$) is inessential and we can simplify the problem to the quadratic tree level superpotential given by (5.2.14). For this superpotential, we will compute the effective glueball superpotential using the formalism reviewed in the previous section. Then, by solving the equation of motion, we study the *on-shell* process of sending poles through the cut, and see whether the poles can pass or not.

Also, on the way, we make an observation on the relation between the exact superpotential and the vacuum expectation value of the tree level superpotential.

5.3.1 The off-shell W_{eff}, $M(z)$ and $T(z)$

First, let us compute the effective glueball superpotential for the quadratic superpotential (5.2.14). The matrix model curve (5.2.7) is related to $W'(z)$ in this

case as

$$y_m^2 = W'(z)^2 + f_0(z) = m_A^2 z^2 - \mu \equiv m_A^2(z^2 - \widetilde{\mu}), \qquad \widetilde{\mu} = \frac{\mu}{m_A^2}. \quad (5.3.1)$$

Let us consider the case with K poles on the first sheet at $q_I = q_f$, for which $r_I = 1$, and with $(N_f - K)$ poles on the second sheet at $\widetilde{q}_I = \widetilde{q}_f$, for which $r_I = 0$ (recall that $(q_f, \widetilde{q}_f)$ is the lift of z_f defined in (5.2.15)). Using the curve (5.3.1) and various formulas summarized in the previous section, one can compute

$$\begin{aligned}
S &= \frac{1}{2\pi i}\oint_A R(z)dz = \frac{m_A\widetilde{\mu}}{4} = \frac{\mu}{4m_A}, \\
\Pi &= 2\int_{\sqrt{\widetilde{\mu}}}^{\Lambda_0} y(z)dz = m_A\Lambda_0^2 - 2S - 2S\log\frac{\Lambda_0^2 m_A}{S}, \\
\Pi_{f,I}^{r_I=0} &= \int_{\widetilde{q}_I}^{\widetilde{\Lambda}_0} y(z)dz = -\int_{q_I}^{\Lambda_0} y(z)dz \\
&= \frac{-m_A\Lambda_0^2}{2} - 2S\log\frac{z_I}{\Lambda_0} \\
&\quad +2S\left[-\log\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{4S}{m_A z_I^2}}\right) + \frac{m_A z_I^2}{4S}\sqrt{1-\frac{4S}{m_A z_I^2}} + \frac{1}{2}\right], \\
\Pi_{f,I}^{r_I=1} &= \int_{q_I}^{\widetilde{\Lambda}_0} y(z)dz = -\int_{\widetilde{q}_I}^{\Lambda_0} y(z)dz \\
&= \frac{-m_A\Lambda_0^2}{2} - 2S\log\frac{z_I}{\Lambda_0} \\
&\quad +2S\left[-\log\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{4S}{m_A z_I^2}}\right) - \frac{m_A z_I^2}{4S}\sqrt{1-\frac{4S}{m_A z_I^2}} + \frac{1}{2}\right],
\end{aligned}$$

where we dropped $\mathcal{O}(1/\Lambda_0)$ terms. We have traded $q_I, \widetilde{q}_I$ for z_I in the square roots, so that the sign convention is such that $\sqrt{1-\frac{4S}{m_A z_I^2}} \sim 1 - \frac{2S}{m_A z_I^2}$ and $\sqrt{z_I^2 - \frac{4S}{m_A}} \sim$

z_I for $|z_I|$ very large. Substituting this into (5.2.12), we obtain

$$\begin{aligned} W_{\rm eff}(S) &= S\left[N_c + \log\left(\frac{m_A^{N_c}\Lambda^{2N_c-N_f}\prod_I z_I}{S^{N_c}}\right)\right] \\ &- \sum_{I,r_I=0} S\left[-\log\left(\frac{1}{2}+\frac{1}{2}\sqrt{1-\frac{4S}{m_A z_I^2}}\right)+\frac{m_A z_I^2}{4S}\left(\sqrt{1-\frac{4S}{m_A z_I^2}}-1\right)+\frac{1}{2}\right] \\ &- \sum_{I,r_I=1} S\left[-\log\left(\frac{1}{2}-\frac{1}{2}\sqrt{1-\frac{4S}{m_A z_I^2}}\right)+\frac{m_A z_I^2}{4S}\left(-\sqrt{1-\frac{4S}{m_A z_I^2}}-1\right)+\frac{1}{2}\right], \end{aligned} \tag{5.3.2}$$

where Λ is the dynamical scale of the corresponding $\mathcal{N}=2$ gauge theory defined in (5.2.13).

Let us compute resolvents also. The resolvent $M(z)$ is an $N_f \times N_f$ matrix. Using (5.2.8), we find that for the I-th eigenvalue

$$M_I(z) = \frac{-R(z)}{(z-q_I)} + \frac{(1-r_I)R(q_I)}{(z-q_I)} + \frac{r_I R(\widetilde{q}_I)}{(z-q_I)}.$$

Expanding $M_I(z)$ around $z=\infty$ we can read off the following vacuum expectation values

$$\begin{aligned} \langle \widetilde{Q}Q\rangle_I &= \frac{m_A}{2}\left[z_I + (2r_I-1)\sqrt{z_I^2-\widetilde{\mu}}\right], \\ \langle \widetilde{Q}\Phi Q\rangle_I &= \frac{m_A}{4}\left[2z_I^2 + 2(2r_I-1)z_I\sqrt{z_I^2-\widetilde{\mu}}-\widetilde{\mu}\right], \\ \left\langle \sum_I -(\widetilde{Q}\Phi Q - z_I\widetilde{Q}Q)\right\rangle &= \frac{N_f m_A\widetilde{\mu}}{4} = N_f S. \end{aligned}$$

There is something worth noting here. One might naively expect that the exact superpotential is simply the vacuum expectation value of the tree level superpotential (5.2.14), as is the case without flavors. However, this naive expectation is wrong! Although we have $\langle W_{\rm tree,fund}\rangle = \langle -\sum_I(\widetilde{Q}\Phi Q - z_I\widetilde{Q}Q)\rangle \neq 0$ if $S\neq 0$, we still have

$$W_{\text{eff,on-shell}} = \left\langle \frac{m_A}{2}\mathrm{Tr}\,\Phi^2\right\rangle \ , \tag{5.3.3}$$

as we will see shortly. The reason for (5.3.3) can be explained by symmetry arguments [93, 77]. Although the tree level part $\langle W_{\text{tree,fund}}\rangle$ for fundamentals is generically nonzero and also contributes to W_{low}, the contribution is precisely canceled by the dynamically generated superpotential W_{dyn}, leaving only the Φ part of W_{tree} [4].

Let us calculate the resolvent $T(z)$ also. For the present example, $c(z) = m_A(N_c - \frac{N_f}{2})$ from (5.2.10). Therefore, using (5.2.9), we obtain the expansion of the resolvent $T(z)$:

$$\begin{aligned} T(z) &= \frac{N_c}{z} + \sum_I \left[\frac{z_I}{2} + \frac{2r_I - 1}{2}\sqrt{z_I^2 - \widetilde{\mu}}\right]\frac{1}{z^2} \\ &+ \left[\frac{\widetilde{\mu}(2N_c - N_f)}{4} + \sum_I \left(\frac{z_I^2}{2} + \frac{2r_I - 1}{2} z_I \sqrt{z_I^2 - \widetilde{\mu}}\right)\right]\frac{1}{z^3} + \cdots \end{aligned} \tag{5.3.4}$$

From this we can read off $\langle \operatorname{Tr}\Phi^n\rangle$. For example, for $K = 0$ we have

$$\begin{aligned} \langle \operatorname{Tr}\Phi\rangle &= \frac{N_f}{2}\left(z_f - \sqrt{z_f^2 - \widetilde{\mu}}\right), \\ \left\langle \operatorname{Tr}\Phi^2\right\rangle &= \frac{N_f}{2} z_f \left(z_f - \sqrt{z_f^2 - \widetilde{\mu}}\right) + \frac{(2N_c - N_f)\widetilde{\mu}}{4}. \end{aligned}$$

5.3.2 The on-shell solution

Now we can use the above off-shell expressions to find the on-shell solution. First we rewrite the superpotential as

$$\begin{aligned} W_{\text{eff}} &= S\left[N_c + \log\left(\frac{\Lambda_1^{3N_c - N_f}}{S^{N_c}}\right)\right] \\ &- (N_f - K)S\left[-\log\left(\frac{z_f}{2} + \frac{1}{2}\sqrt{z_f^2 - \frac{4S}{m_A}}\right) + \frac{m_A z_f}{4S}\left(\sqrt{z_f^2 - \frac{4S}{m_A}} - z_f\right) + \frac{1}{2}\right] \\ &- KS\left[-\log\left(\frac{z_f}{2} - \frac{1}{2}\sqrt{z_f^2 - \frac{4S}{m_A}}\right) + \frac{m_A z_f}{4S}\left(-\sqrt{z_f^2 - \frac{4S}{m_A}} - z_f\right) + \frac{1}{2}\right], \end{aligned} \tag{5.3.5}$$

[4] We would like to thank K. Intriligator for explaining this point to us.

where $\Lambda_1^{3N_c-N_f} \equiv m_A^{N_c}\Lambda^{2N_c-N_f}$. We have set all z_I to be z_f, i.e., all masses are the same, as in (5.2.15). From this we obtain[5]

$$0 = \log\left(\frac{\Lambda_1^{3N_c-N_f}}{S^{N_c}}\right) + K\log\left(\frac{z_f - \sqrt{z_f^2 - \frac{4S}{m_A}}}{2}\right) + (N_f - K)\log\left(\frac{z_f + \sqrt{z_f^2 - \frac{4S}{m_A}}}{2}\right), \tag{5.3.6}$$

or [6]

$$0 = \log\left(\widehat{S}^{-N_c}\right) + K\log\left(\frac{\widehat{z}_f - \sqrt{\widehat{z}_f^2 - 4\widehat{S}}}{2}\right) + (N_f - K)\log\left(\frac{\widehat{z}_f + \sqrt{\widehat{z}_f^2 - 4\widehat{S}}}{2}\right), \tag{5.3.7}$$

where we have defined dimensionless quantities $\widehat{S} = \frac{S}{m_A\Lambda^2} = \frac{\widetilde{\mu}}{4\Lambda^2}$ and $\widehat{z}_f = \frac{z_f}{\Lambda}$. Note that using these massless quantities the cut is from along the interval $[-2\sqrt{\widehat{S}}, 2\sqrt{\widehat{S}}]$.

Using (5.3.6) or (5.3.7) it is easy to show that

$$W_{\text{eff,on-shell}} = \left(N_c - \frac{N_f}{2}\right)S + \frac{m_A z_f^2}{4}\left[N_f + (2K - N_f)\sqrt{1 - \frac{4S}{m_A z_f^2}}\right].$$

[5]It is easy to check that the equation (5.3.6) with parameters (N_c, N_f, K) is the same as the one with parameters $(N_c - r, N_f - 2r, K - r)$. Also from the expression (5.3.4) it is straightforward to see we have $\left\langle \mathrm{Tr}\Phi^2\right\rangle_{N_c,N_f,K} = \left\langle \mathrm{Tr}\Phi^2\right\rangle_{N_c-r,N_f-2r,K-r} + rz_I^2$. All of these facts are the result of the "addition map" observed in [81]. Furthermore, one can show that both (5.3.6) and (5.3.4) for $K = 0$ are exactly the same as the one given by the strong coupling analysis in [92] and the weak coupling analysis in [94, 95].

[6]As we mentioned before, for $m_f \neq 0$ the allowed Higgs branch requires $K \leq N_f$ but the strong coupling analysis gives $K \leq N_f/2$. The resolution for that puzzle is that if $K > N_f/2$, it is given by $(N_f - K)$-th branch of the curve. Since the same r-th branch of the curve gives both r and $(N_f - r)$ Higgs branches, we expect that r and $(N_f - r)$ Higgs branches are related. This relation is given by $\widetilde{S} = Sb^2$, $\widetilde{z}_f = z_f b$ with $b = \frac{m_A\Lambda^2}{S}$. It can be shown that with the above relation, the equation of motion of S for the K-th branch is changed to the equation of motion of $\widetilde{S}$ for the $(N_f - K)$-th branch.

Also by expanding $T(z)$ in the present case, just as we did in (5.3.4), we can read off

$$\left\langle \frac{m_A}{2} \mathrm{Tr}\, \Phi^2 \right\rangle = \left(N_c - \frac{N_f}{2} \right) S + \frac{m_A z_f^2}{4} \left[N_f + (2K - N_f) \sqrt{1 - \frac{4S}{m_A z_f^2}} \right],$$

which gives us the relation

$$\left\langle \frac{m_A}{2} \mathrm{Tr}\, \Phi^2 \right\rangle = W_{\text{eff,on-shell}}$$

as we promised in (5.3.3).

Equation (5.3.6) is hard to solve. But if we want just to discuss whether the cut closes up when we bring $z_f \to 0$, we can set $K = 0$ [7], for which (5.3.6) reduces to

$$z_f = \omega_{N_f}^{r} \widehat{S}^{\frac{N_c}{N_f}} + \omega_{N_f}^{-r} \widehat{S}^{\frac{N_f - N_c}{N_f}}. \tag{5.3.8}$$

Here, ω_{N_f} is the N_f-th root of unity, $\omega_{N_f} = e^{2\pi i/N_f}$, and $r = 0, 1, \ldots, (N_f - 1)$ corresponds to different branches of solutions. It is also amusing to note that above solution has the Seiberg duality [96] where electric theory with (N_c, N_f) is mapped to a magnetic theory with $(N_f - N_c, N_f)$.

With these preparations, we can start to discuss the on-shell process of passing poles through the cut from the second sheet.

5.3.3 Passing poles through a branch cut

Consider moving N_f poles on top of each other at infinity on the second sheet toward the cut along a line passing through the origin and making an angle of θ

[7] If $K \neq 0$, we will have $U(N_c) \to U(K) \times U(N_c - K)$ and the problem reduces to of $U(N_c - K)$ with $(N_f - 2K)$ flavors in the 0-th branch.

with the real axis. Namely, take[8]

$$z(\widetilde{q}_f) = z(q_f) = pe^{i\theta}, \qquad p, \theta \in \mathbb{R}, \tag{5.3.9}$$

and change p from $p = \infty$ to $p = -\infty$ (see Fig. 5.1). This equation (5.3.9) needs some explanation. Remember that $z_f = z(q_f) = z(\widetilde{q}_f)$ denotes the projection from the Riemann surface Σ (Eq. (5.3.1)) to the z-plane. For each point z on the z-plane, there are two corresponding points: q on the first sheet and $\widetilde{q}$ on the second sheet. Although we are starting with poles at $\widetilde{q}_f$ on the second sheet, we do not know in advance if the poles will pass through the cut and end up on the first sheet, or it will remain on the second sheet. Therefore we cannot specify which sheet the poles are on, and that is why we used $z(q_f)$, $z(\widetilde{q}_f)$ in (5.3.9), instead of q_f or $\widetilde{q}_f$.

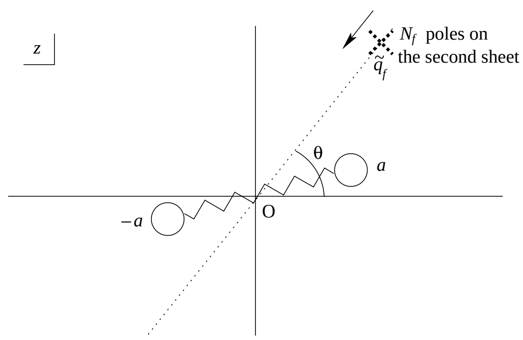


Figure 5.1: *A process in which N_f poles at $\widetilde{q}_f$ on the second sheet far away from the cut approach the branch cut on the double sheeted z-plane, along a line which goes through the origin and makes an angle θ with the real z axis. The "$\times$" with dotted lines denotes the poles on the second sheet, moving in the direction of the arrow. The two branch points $\pm a$ are connected by the branch cut, which is denoted by a zigzag.*

[8]Throughout this subsection, we will use the dimensionless quantities $\widehat{z}_f$, $\widehat{S}$, etc. and omit the hats on them to avoid clutter, unless otherwise mentioned.

Below, we study the solution to the equation of motion (5.3.8), changing $p \in \mathbb{R}$ from $p = \infty$ to $p = -\infty$. By redefining z_f, S by $z_f \to z_f e^{2\pi i r/(N_f - 2N_c)}$, $S \to S e^{4\pi i r/(N_f - 2N_c)}$, we can bring (5.3.8) to the following form:

$$z_f = p e^{i\theta} = S^t + S^{1-t}, \tag{5.3.10}$$

where

$$\frac{N_c}{N_f} \equiv t.$$

Henceforth we will use (5.3.10). Because z_f as well as S is complex, the position of the branch points (namely, the ends of the cut), $\pm a$, where

$$a \equiv \sqrt{4S}$$

is also complex, which means that in general the cut makes some finite angle with the real axis, as shown in Fig. 5.1.

- $\boldsymbol{N_f = N_c}$

As the simplest example, let us first consider the $N_f = N_c$ (i.e., $t = 1$) case. We will see that the poles barely pass through the cut but get soon bounced back to the second sheet.

The equation of motion (5.3.10) is, in this case,

$$z_f = p e^{i\theta} = S + 1. \tag{5.3.11}$$

Therefore, as we change p, the position of the branch points changes according to

$$a = \sqrt{4S} = 2\sqrt{z_f - 1} = 2\sqrt{p e^{i\theta} - 1}. \tag{5.3.12}$$

Let us look closely at the process, step by step. The point is that transition between the first and the second sheet can happen only when the cut becomes parallel to the incident direction of the poles, or when the poles pass through the origin.

(1) $p \simeq +\infty$, on the second sheet:

In this case, we can approximate the right hand side of (5.3.12) as

$$\begin{aligned} a &= 2p^{\frac{1}{2}} e^{\frac{i\theta}{2}} (1 - p^{-1} e^{-i\theta})^{\frac{1}{2}} \\ &\simeq 2p^{\frac{1}{2}} e^{\frac{i\theta}{2}} e^{-\frac{1}{2} p^{-1} e^{-i\theta}} = 2p^{\frac{1}{2}} e^{-\frac{1}{2} p^{-1} \cos\theta} \; e^{i(\frac{\theta}{2} + \frac{1}{2} p^{-1} \sin\theta)}. \end{aligned}$$

Therefore, when the poles are far away, the angle between the cut and the real axis is approximately $\frac{\theta}{2} > 0$ (we assume $0 < \theta < \frac{\pi}{2}$). Furthermore, as the poles approach (p becomes smaller), the cut shrinks (because of $p^{\frac{1}{2}}$) and rotates counterclockwise (because of $e^{\frac{i}{2} p^{-1} \sin\theta}$). This corresponds to Fig. 5.2a.

(2) Because the cut is rotating counterclockwise, as the poles approach, the cut will eventually become parallel to the incident direction, at some point. This happens when

$$a^2 = 4(pe^{i\theta} - 1) = 4\left[(p\cos\theta - 1) + ip\sin\theta\right] \propto e^{2i\theta} = \cos 2\theta + i \sin 2\theta.$$

By simple algebra, one obtains

$$p = 2\cos\theta, \qquad a = 2e^{i\theta}. \tag{5.3.13}$$

Note that this is the only solution; the cut becomes parallel to the incident direction only once. Because $0 < p < |a| = 2$ (we are assuming $0 < \theta < \pi/2$), by the time the cut becomes parallel to the incident direction, the poles have come inside of the interval $[-2e^{i\theta}, 2e^{i\theta}]$, along which the cut

extends when it is parallel to the incident direction. This implies that the poles cross [9] the cut at this point, and enter into the first sheet. Fig. 5.2b shows the situation when this transition is about to happen. In Fig. 5.2c, the poles are just crossing the cut. Fig. 5.2d corresponds to the situation just after the transition happened; the poles have passed the cut and are now proceeding on the first sheet.

Here we implicitly assumed that the cut is still rotating counterclockwise with a finite angular velocity, but this can be shown by expanding a around (5.3.13) as $p = 2\cos\theta + \Delta p$. A short computation shows

$$a \simeq 2e^{\frac{1}{2}\Delta p\,\cos\theta} e^{i\theta - \frac{i}{2}\Delta p\,\sin\theta}$$

which implies that the cut shrinks and rotates counterclockwise if we move the poles to the left (Δp decreases).

(3) $p \simeq 0$:

If the poles proceed on the real line further, it eventually reaches the origin $p = 0$. By expanding (5.3.12) around $p = 0$, one obtains

$$a = 2(e^{\pi i} + pe^{i\theta})^{\frac{1}{2}} \simeq 2e^{-\frac{1}{2}p\cos\theta} e^{i(\frac{\pi}{2} - \frac{1}{2}p\sin\theta)}. \tag{5.3.14}$$

Therefore the cut has a finite size ($|a| = 2$) at $p = 0$ and along the imaginary axis, still rotating counterclockwise, but now expanding. Because the cut goes through the origin, the poles pass through the cut again and comes back onto the second sheet (Fig. 5.2e).

(4) $p \simeq -\infty$, on the second sheet:

If the poles have gone far past the cut so that $p < 0$, $|p| \gg 1$, we can

[9] As mentioned in footnote 2, there is no real physical meaning to the position of the cut itself.

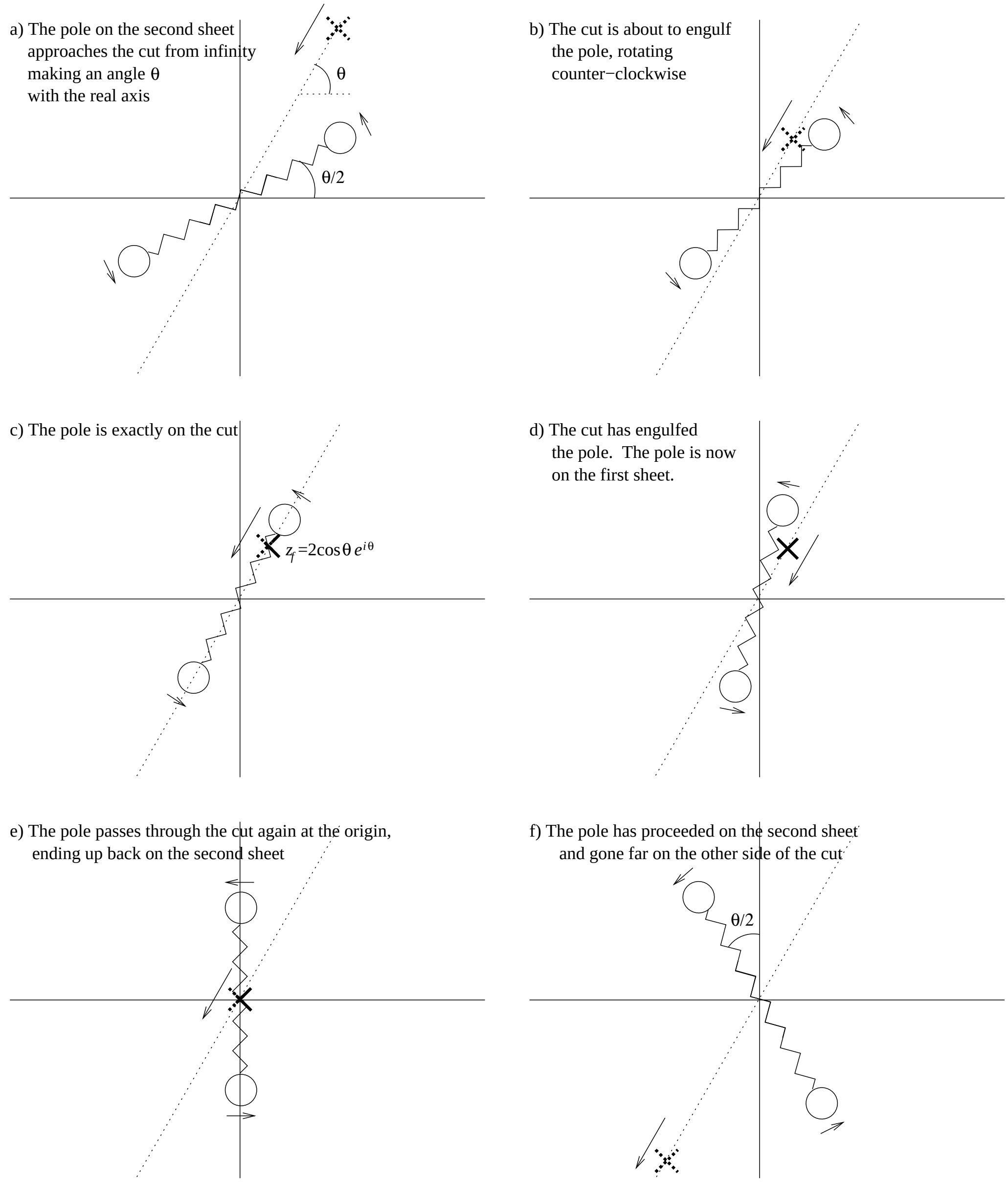


Figure 5.2: *Six configurations of the branch cut and the poles. The poles are depicted by "×" and moving along a line at an angle θ with the real axis, as the arrow on it indicates. The "×" in solid (dotted) lines denotes poles on the first (second) sheet. The branch cut is rotating counterclockwise (as the arrows on its sides indicate), changing its length.*

approximate (5.3.12), as before, as

$$a = 2|p|^{\frac{1}{2}} e^{\frac{i}{2}(\theta+\pi)} (1 - p^{-1} e^{-i\theta})^{\frac{1}{2}} \simeq 2|p|^{\frac{1}{2}} e^{\frac{1}{2}|p|^{-1}\cos\theta}\ e^{i(\frac{\pi}{2}+\frac{\theta}{2}-\frac{1}{2}|p|^{-1}\sin\theta)}.$$

Therefore, as the poles go away, the cut expands and rotates counterclockwise. The angle between the cut and the real axis asymptotes to $(\frac{\pi}{2}+\frac{\theta}{2})$ (Fig. 5.2f).

In the above we assumed that $0 < \theta < \frac{\pi}{2}$. If $\frac{\pi}{2} < \theta < \pi$, the only difference is that the order of steps (2) and (3) are exchanged. If $\theta < 0$, the cut rotates clockwise instead of counterclockwise.

When is the cut shortest in this whole process? From (5.3.12), one easily obtains

$$|a| = 2[(p-\cos\theta)^2 + \sin^2\theta]^{1/4} \geq 2|\sin\theta|^{1/2}. \tag{5.3.15}$$

Therefore, when the poles are at $z_f = pe^{i\theta} = \cos\theta\, e^{i\theta}$, which is between the steps (2) and (3) above, the cut becomes shortest. In particular, in the limit $\theta \to 0$ or $\theta \to \pm\pi$, the cut completely closes up instantaneously. These correspond to configurations with either a horizontal cut with poles colliding sideways, or a vertical cut with poles colliding from right above or from right below. Actually the existence of the $S = 0$ solution is easy to see in (5.3.11): it is just $z_f = 1, S = 0$.

Summary: for $N_f = N_c$, when one moves poles on the second sheet from infinity along a line toward a cut, poles pass through the cut onto the first sheet and move away from the cut by a short distance. Then poles are bounced back to the second sheet again. Therefore, one can never move poles far away from the cut on the first sheet. During the process, in certain situations, the cut completely closes up.

- $\boldsymbol{N_f \neq N_c}$

Now let us consider a more general case with $N_f \neq N_c$. We again consider a situation where poles on the second sheet approach a cut. This time we will be brief and sketchy, because a detailed analysis such as the one we did for the $N_f = N_c$ case would be rather lengthy due to the existence of multiple branches, and would not be very illuminating.

First, let us ask how we can see whether poles are on the first sheet or on the second sheet, from the behavior of S versus p. Because this is not apparent in the equation of motion of the form (5.3.10), let us go back to

$$0 = \frac{\partial W_{\text{eff}}}{\partial S} \propto \log S^{-N_c} + N_f \log \frac{z_f \mp \sqrt{z_f^2 - 4S}}{2} \quad \Longrightarrow \quad S^t = \frac{z_f \mp \sqrt{z_f^2 - 4S}}{2}. \tag{5.3.16}$$

which led to the equation (5.3.10). Here the "$-$" ("$+$") sign corresponds to q_f on the first (second) sheet. For $|z_f|^2 \gg |4S|$, the square root can be approximated as $\sqrt{z_f^2 - 4S} = z_f(1-4S/z_f^2)^{1/2} \simeq z_f(1-2S/z_f^2)$ (our sign convention was discussed above (5.3.2)). Therefore (5.3.16) is, on the first sheet,

$$S^t \simeq \frac{z_f - z_f(1-2S/z_f^2)}{2} = \frac{S}{z_f} \qquad \Longrightarrow \qquad z_f \simeq S^{1-t}, \tag{5.3.17}$$

while on the second sheet

$$S^t \simeq \frac{z_f + z_f(1-2S/z_f^2)}{2} \simeq z_f \qquad \Longrightarrow \qquad z_f \simeq S^t. \tag{5.3.18}$$

Now let us solve (5.3.10) for $|z_f| \gg 1$. By carefully comparing the magnitude

of the two terms in (5.3.10), one obtains

$$\begin{array}{llll} N_f < N_c & (1 < t) & \rightarrow & \begin{cases} z_f \simeq S^t & \rightarrow \quad |S| \simeq |p|^{\frac{1}{t}}, \qquad |S| \gg 1 \\ z_f \simeq S^{1-t} & \rightarrow \quad |S| \simeq |p|^{-\frac{1}{t-1}}, \quad |S| \ll 1 \end{cases} \\ N_c < N_f < 2N_c & (\frac{1}{2} < t < 1) & \rightarrow & \quad z_f \simeq S^t \quad\quad \rightarrow \quad |S| \simeq |p|^{\frac{1}{t}}, \qquad |S| \gg 1 \\ 2N_c < N_f & (0 < t < \frac{1}{2}) & \rightarrow & \quad z_f \simeq S^{1-t} \quad \rightarrow \quad |S| \simeq |p|^{\frac{1}{1-t}}, \qquad |S| \gg 1 \end{array} \tag{5.3.19}$$

It is easy to show that $|z_f|^2 \gg |4S|$ in all cases. So by using (5.3.17), (5.3.18), we conclude that the first and the third lines in (5.3.19) correspond to poles on the second sheet, while the second and the last lines correspond to poles on the first sheet. This implies that, only for $N_f < N_c$, poles on the second sheet can pass through the cut all the way and go infinitely far away on the first sheet from a cut, as we will see explicitly in the examples below. For $N_c < N_f < 2N_c$, if one tries to pass poles through a cut, then either poles will be bounced back to the second sheet, or the cut closes up before the poles reach it. For $N_f > 2N_c$, there is no solution corresponding to poles moving toward a cut from infinity on the second sheet. This should be related to the fact that in this case glueball S is not a good IR field. Therefore, this one cut model is not applicable for $N_f > 2N_c$.

To argue that these statements are true, rather than doing an analysis similar to the one we did for the $N_f = N_c$ case, we will present some explicit solutions for some specific values of $t = \frac{N_c}{N_f}$ and θ, and argue general features. Before looking at explicit solutions, note that there are multiple solutions to Eq. (5.3.10) which can be written as

$$z_f = (S^{\frac{1}{N_f}})^{N_c} + (S^{\frac{1}{N_f}})^{N_f - N_c}. \tag{5.3.20}$$

From the degree of this equation, one sees that the number of the solutions to

(5.3.20) is:

$$\begin{aligned}
N_f \leq N_c \quad &\Longrightarrow \quad 2N_c - N_f \text{ solutions,} \\
N_c \leq N_f \leq 2N_c \quad &\Longrightarrow \quad N_c \text{ solutions,} \\
2N_c \leq N_f \quad &\Longrightarrow \quad N_f - N_c \text{ solutions.}
\end{aligned}$$

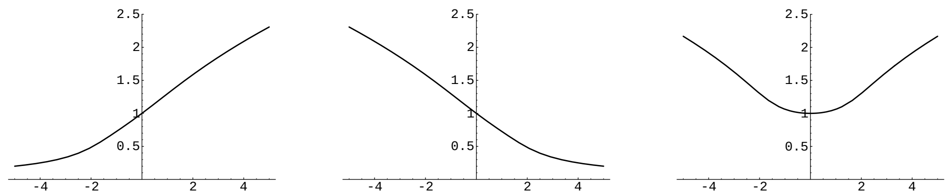
Therefore, if we solve the equation of motion (5.3.10), in general we expect multiple branches of solutions. We do not have to consider the N_f branches of the root $S^{\frac{1}{N_f}}$, because it is taken care of by the phase rotation we did above (5.3.10).

Now let us look at explicit solutions for $U(2)$ example. For $N_f = \frac{1}{2}N_c$ ($t = 2$), the solution to the equation of motion (5.3.10) is

$$S = -\frac{1}{3}\left(\frac{27 + \sqrt{729 - 108z_f^3}}{2}\right)^{1/3} - \left(\frac{27 + \sqrt{729 - 108z_f^3}}{2}\right)^{-1/3},$$

where three branches of the cubic root are implied. In Fig. 5.3 we plotted $|S|$ versus p for these branches, for a randomly chosen value of the angle of incidence, $\theta = \pi/6$. Even if one changes θ, there are always three branches whose general shapes are similar to the ones in Fig. 5.3. These three branches changes into one another when θ is changed by $2\pi/3$. One can easily see which branch corresponds to what kind of processes, by the fact that on the first sheet $|S| \ll 1$ as $|p| \to \infty$, while on the second sheet $|S| \gg 1$ as $|p| \to \infty$. The three branches correspond to the process in which: i) poles go from the second sheet to the first sheet through the cut, without any obstruction, ii) poles go from the first sheet to the second sheet (this is not the process we are interested in), and iii) poles coming from the second sheet get reflected back to the second sheet. Note that the cut has never closed in all cases, because $|S|$ is always nonvanishing.

Similarly, for $N_f = \frac{3}{2}N_c$ ($t = \frac{2}{3}$), the solution to the equation of motion



branch i) branch ii) branch iii)

Figure 5.3: *The graph of* $|S|$ *versus* p *for* $N_f = \frac{1}{2}N_c$ $(t = 2)$, $\theta = \pi/6$. *The vertical axis is* $|S|$ *and the horizontal axis is* p. *Although we are showing just the* $\theta = \pi/6$ *case, there are similar looking three branches for any value of* θ, *which change into one another when* θ *is changed by* $2\pi/3$.

(5.3.10) is

$$S = \frac{1}{2}\left[-3z_f - 1 \pm (z_f + 1)\sqrt{4z_f + 1}\right].$$

This time there are two branches, which change into each other when the θ is changed by π. We plotted $|S|$ versus p for $\theta = \pi/2$ in Fig. 5.4. It shows two possibilities: i) poles coming from the second sheet get reflected back to the second sheet, for which $|S| \neq 0$ as $p \to 0$, ii) the cut closes up before poles passes through it, for which $|S| \to 0$ as $p \to 0$.

The $N_f = 3N_c$ $(t = \frac{1}{3})$ case is also described by the same Fig. 5.4. However, as we discussed below (5.3.19), it does not correspond to a process of poles approaching the cut from infinity on the second sheet; it corresponds to poles on the first sheet and we cannot give any physical interpretation to it.

These demonstrate the following general features:

- For $N_f < N_c$, one can move poles at infinity on the second sheet through a cut all the way to infinity on the first sheet without obstruction, if one

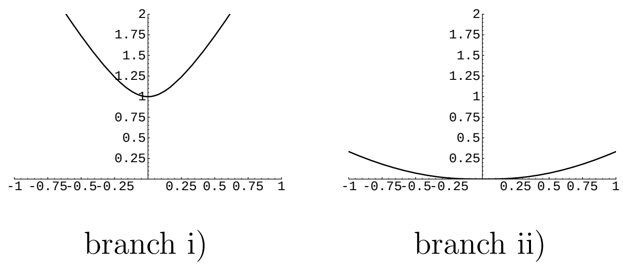


Figure 5.4: *The graph of* $|S|$ *versus* p *for* $N_f = \frac{3}{2}N_c$ $(t = \frac{2}{3})$ *or* $N_f = 3N_c$ $(t = \frac{1}{3})$, *for* $\theta = \pi/2$. *The two values of* t *give the same graph. The vertical axis is* $|S|$ *and the horizontal axis is* p. *Although we are showing just the* $\theta = \pi/2$ *case, there are similar looking two branches for any value of* θ, *which change into each other when* θ *is changed by* π,

chooses the incident angle appropriately. If the angle is not chosen appropriately, the poles will be bounced back to the second sheet.

- For $N_c \leq N_f < 2N_c$, one cannot move poles at infinity on the second sheet through a cut all the way to infinity on first sheet. If one tries to, either i) the cut rotates and sends the poles back to the second sheet, or ii) the cut closes up before the poles reach it.

- For $2N_c < N_f$, the one cut model does not apply directly.

$N_f = 2N_c$ is an exceptional case, for which the equation of motion (5.3.10) becomes

$$z_f = 2S^{1/2}. \tag{5.3.21}$$

Therefore $|S| \to 0$ as $z_f \to 0$, and the cut always closes before the poles reach it.

Subtlety in $S = 0$ solutions

If $p = 0$, or equivalently if $z_f = 0$, there is a subtle, but important point we overlooked in the above arguments. For $z_f = 0$, the superpotential (5.3.5) becomes

$$\begin{aligned} W_{\text{eff}} &= S\left[\left(N_c - \frac{N_f}{2}\right) + \log\left(\frac{(-1)^{N_f/2} m_A^{N_c - N_f/2} \Lambda^{2N_c - N_f}}{S^{N_c - N_f/2}}\right)\right] \\ &= S\left[\left(N_c - \frac{N_f}{2}\right) + \log\left(\frac{(-1)^{-N_f/2} m_A^{N_c - N_f/2} \Lambda_0^{2N_c - N_f}}{S^{N_c - N_f/2}}\right)\right] + 2\pi i \tau_0 S \\ &= \left(N_c - \frac{N_f}{2}\right) S\left[1 + \log\left(\frac{\widetilde{\Lambda}_0^3}{S}\right)\right] + 2\pi i \tau_0 S \end{aligned} \tag{5.3.22}$$

with $\widetilde{\Lambda}_0^{3(N_c - N_f/2)} \equiv (-1)^{-N_f/2} m_A^{N_c - N_f/2} \Lambda_0^{2N_c - N_f}$. Only from here to (5.3.23), S means the dimensionful quantity ($S = m_A \Lambda^2 \widehat{S}$; see below (5.3.7)). In addition, in the second line of (5.3.22), we rewrite the renormalized scale Λ in terms of the bare scale Λ_0 and the bare coupling τ_0 using the relation (5.2.13). In our case, $B_L = (-1)^{N_f}$. The equation of motion derived from (5.3.22) is

$$\left(N_c - \frac{N_f}{2}\right) \log\left(\frac{\widetilde{\Lambda}_0^3}{S}\right) + 2\pi i \tau_0 = 0$$

and the solution is

$$S = \begin{cases} \widetilde{\Lambda}_0^3 \, e^{\frac{2\pi i \tau_0}{N_c - N_f/2}} & N_f \neq 2N_c \ , \\ \text{no solution} & N_f = 2N_c \ . \end{cases} \tag{5.3.23}$$

For $N_f < N_c$, (5.3.23) is consistent with the fact that the $|S|$ versus p graphs in Fig. 5.3 all go through the point $(p, |S|) = (0, 1)$ (now S means the dimensionless quantity). Also for $N_f = N_c$, (5.3.23) is consistent with the result (5.3.14) ($|a| = 2$, so $|S| = 1$). On the other hand, for $N_f > N_c$, (5.3.23) implies that we should exclude the origin $(p, |S|) = (0, 0)$ from the $|S|$-p graphs in Fig. 5.4, which is the only $S = 0$ solution (this includes the $N_f = 2N_c$ case (5.3.21)). [10]

[10] One may think that if one uses the first line of (5.3.22), then for $N_f = 2N_c$, $W_{\text{eff}} =$

Therefore, the above analysis seems to indicate that, for $N_f > N_c$, the $S = 0$ solution at $p = 0$, or equivalently $z_f = 0$ is an exceptional case and should be excluded. On the other hand, as can be checked easily, gauge theory analysis based on the factorization method shows that there is an $S = 0$ solution in the baryonic branch. Thus we face the problem of whether the baryonic $S = 0$ branch for $N_f > N_c$ can be described in matrix model, as alluded to in the previous discussions.

Note that, there is also an $S = 0$ solution for $N_c = N_f$ in certain situations, as discussed below (5.3.15). For this solution, which is in the non-baryonic branch, there is no subtlety in the equation of motion such as (5.3.23), and it appears to be a real on-shell solution. This will be discussed further below.

5.3.4 Generalization of IKRSV

In the above and in subsection 5.2.4, we argued that for $N_f \geq N_c$ the $S = 0$ solutions are real, on-shell solutions based on the factorization analysis. More accurately, there are two cases with $S = 0$: the one in the maximal non-baryonic branch with $N_f \geq 2N_c$ and the other one in the baryonic branch with $N_c \leq N_f < 2N_c$. The case of non-baryonic branch cannot be discussed in the one cut model, which is applicable only to $N_f < 2N_c$. On the other hand, the baryonic one did show up in the previous subsection, but we just saw above that those solutions should be excluded by the matrix model analysis. What is happening? Is it impossible to describe the baryonic branch in matrix model?

Recall that the glueball field S has to do with the strongly coupled dynamics of

$S \log[(-1)^{N_f/2}]$ and there are solutions for some N_f. However, the glueball superpotential that string theory predicts [10, 80] is the third line of (5.3.22) which is in terms of the bare quantities Λ_0 and τ_0. If $N_f = 2N_c$, then the log term vanishes and one cannot define a new scale Λ as we did in (5.2.13) to absorb the linear term $2\pi i \tau_0 S$.

$U(N_c)$ theory. That $S = 0$ in those solutions means that there is no strongly coupled dynamics any more, namely the $U(N_c)$ group has broken down completely. The only mechanism for that to happen is by condensation of a massless charged particle which makes the $U(1)$ photon of the $U(N_c)$ group massive. Therefore, in order to make $S = 0$ a solution, we should incorporate such an extra massless degree of freedom, which is clearly missing in the description of the system in terms only of the glueball S. This extra degree of freedom should exist even in the $N_f = N_c$ case where $S = 0$ really is an on-shell solution as discussed below (5.3.15); we just could not directly see the degree of freedom in this case.

The analysis of [87] hints on what this extra massless degree of freedom should be in the matrix model / string theory context. Note that, the superpotential (5.3.22) is of exactly the same form as Eq. (4.5) of [87], if we interpret $N_c - N_f/2 \equiv \widehat{N}$ as the amount of the net RR 3-form fluxes. In [87] it was argued that, if the net RR flux $\widehat{N}$ vanishes, one should take into account an extra degree of freedom corresponding to D3-branes wrapping the blown up S^3 in the Calabi–Yau geometry [60], and condensation of this extra degree of freedom indeed makes $S = 0$ a solution to the equation of motion. The form of the superpotential (5.3.22) strongly suggests that the same mechanism is at work for $N_f = 2N_c$ in the $r = N_c$ non-baryonic branch; condensation of the D3-brane makes $S = 0$ a solution. Furthermore, as discussed in [87], for $N_f > 2N_c$ the glueball S is not a good variable and should be set to zero. A concise way of summarizing this conclusion is: if the generalized dual Coxeter number $h = N_c - N_f/2$ is zero or negative, we should set S to zero in the $r = N_c$ non-baryonic branch.

However, this is not the whole story, as we have discussed in subsection 5.2.4. As we saw above, we need some extra physics also for $N_c \leq N_f < 2N_c$ in order to explain the matrix model result in the baryonic branch. We argue below that

this extra degree of freedom at least in the $N_c < N_f < 2N_c$ case should also be the D3-brane wrapping S^3 which shrinks to zero when the glueball goes to zero: $S \to 0$.

The original argument of [87] is not directly applicable for $N_f < 2N_c$ because there are nonzero RR fluxes penetrating such a D3-brane ($\widehat{N} \neq 0$). These RR fluxes induce fundamental string charge on the D3-brane. Because the D3-brane is compact, there is no place for the flux to end on (note that this flux is not the RR one but the one associated with the fundamental string charge). Hence it should emanate some number of fundamental strings. If there are no flavors, there is no place for such fundamental strings to end on, so they should extend to infinity. This fact led to the conclusion of [87] that the D3-brane wrapping S^3 is infinitely massive and not relevant unless $\widehat{N} = 0$.

However, in our situation, there are places for the fundamental strings to end on — noncompact D5-branes which give rise to flavors [10, 80]. In particular, precisely in the $z_f = 0$ case, where we have $S = 0$ solutions for $N_c < N_f < 2N_c$, the D3-brane wrapping S^3 intersects the noncompact D5-branes in the $S \to 0$ limit, hence the 3-5 strings stretching between them are massless. Therefore the D3-brane with these fundamental strings on it is massless and should be included in the low energy description. It is well known [68] that such a D-brane with fundamental strings ending on it can be interpreted as baryons in gauge theory.[11] Condensation of this baryon degree of freedom should make $S = 0$ a solution, making the photon massive and breaking the $U(N_c)$ down to $U(0)$. The precise form of the superpotential for this extra degree of freedom must be more complicated than the one proposed in [87] for the case without flavors.

[11] That the D3-brane wrapping S^3 cannot exist for $N_f < N_c$ can probably be explained along the same line as [68], by showing that those 3-5 strings are fermionic. Also, note that the gauge group here is $U(N_c)$, not $SU(N_c)$ as in [68], hence the "baryon" is charged under the $U(1)$.

All these analyses tell us the following prescription:

$$\text{\textit{Using the floating mass condition that all N_f poles are on top of one branch point[12] on the Riemann surface, we will have an $S=0$ solution for $N_f \geq 2N_c$. For $N_c < N_f < 2N_c$ there are two solutions: one with $S=0$ in the baryonic branch and one with $S \neq 0$ in the non-baryonic branch. In multi-cut cases, this applies to each cut by replacing N_c, S with the corresponding $N_{c,i}$, S_i for the cut.}} \tag{5.3.24}$$

In the next section we will discuss the condition we have used in above prescription. Also by explicit examples, we will demonstrate that when the gauge theory has a solution with closed cuts ($S_i = 0$), one can reproduce its superpotential in matrix model by setting the corresponding glueballs S_i to zero by hand.

5.4 Two cut model—cubic tree level superpotential

Now, let us move on to $U(N_c)$ theory with cubic tree level superpotential, where we have two cuts. We will demonstrate that for each closed cut we can set $S = 0$ by hand to reproduce the correct gauge theory superpotential using matrix model.

Specifically, we take the tree level superpotential to be

$$\begin{aligned} W_{\text{tree}} &= \mathrm{Tr}[W(\Phi)] - \sum_{I=1}^{N_f} \widetilde{Q}_I(\Phi - z_f)Q^I, \\ W(z) &= \frac{g}{3}z^3 + \frac{m}{2}z^2, \qquad W'(z) = gz\left(z + \frac{m}{g}\right) \equiv g(z-a_1)(z-a_2). \end{aligned} \tag{5.4.1}$$

Here we wrote down $W(z)$ in terms of $g_2 = g$, $g_1 = m$ for definiteness, but mostly we will work with the last expression in terms of g, $a_{1,2}$. The general

[12] This condition will not work for the $N_f = N_{c,i}$ case.

breaking pattern in the pseudo-confining phase is $U(N_c) \to U(N_{c,1}) \times U(N_{c,2})$, $N_{c,1}+N_{c,2} = N_c$, $N_{c,i} > 0$. In the quantum theory, the critical points at a_1 and a_2 blow up into cuts along the intervals $[a_1^-, a_1^+]$ and $[a_2^-, a_2^+]$, respectively. Namely, we end up with the matrix model curve (5.2.7), which in this case is

$$y_m^2 = W'(z)^2 + f_1(z) = g^2(z-a_1^-)(z-a_1^+)(z-a_2^-)(z-a_2^+). \tag{5.4.2}$$

We will call the cuts along $[a_1^-, a_1^+]$ and $[a_2^-, a_2^+]$ respectively the "first cut" and the "second cut" henceforth. One important difference from the quadratic case is that, we can study a process where $N_f \geq 2N_{c,i}$ flavor poles are near the i-th cut in the cubic case.

As we have mentioned, our concern is whether the cut is closed or not. Also from the experiences in the factorization it can be seen that for $N_f > N_{c,i}$, when closed cut is produced, the closed cut and the poles are on top of each other [13]. With all these considerations we take the following condition to constrain the position of the poles:[14]

$$z_f = a_1^-. \tag{5.4.3}$$

If there are $S_1 = 0$ solutions in which the closed cut and the poles are on top of each other, then all such solutions can be found by solving the factorization problem under the constraint (5.4.3), since for such solutions $z = a_1^- = a_1^+$ obviously. One could impose a further condition $S_1 = 0$, or equivalently $a_1^- = a_1^+$ if one wants just closed cut solutions, but we would like to know that there also are solutions with $S_1 \neq 0$ for $N_f < 2N_{c,1}$, so we do not do that.

[13] We do not discuss the $N_f = N_{c,i}$ case where closed cut and poles are not at the same point. However because the $S = 0$ solution in this case is an on-shell solution, we can reproduce the gauge theory result in matrix model without setting $S = 0$ by hand.

[14] We could choose $z_f = a_1^+$ or $z_f = a_2^\pm$ instead of (5.4.3), but the result should be all the same, so we take (5.4.3) without loss of generality.

To summarize, what we are going to do below is: first we explicitly solve the factorization problem under the constraint (5.4.3), and confirm that the $S_1 = 0$ solution exists when $N_f > N_{c,i}$. Then, we reproduce the gauge theory superpotential in matrix model by setting $S_1 = 0$ by hand.

Before plunging into that, we must discuss one aspect of the constraint (5.4.3) and the r-branches, in order to understand the result of the factorization method. If one solves the factorization equation for a given flavor mass $m_f = -z_f$ (without imposing the constraint (5.4.3)), then in general one will find multiple r-branches labeled by an integer K with range $0 \le K \le \min(N_c, [\frac{N_f}{2}])$ (see Eq. (5.2.17)). This is related to the fact that the factorization method cannot distinguish between the poles on the first sheet and the ones on the second sheet. The r-branch labeled by K corresponds to distributing $N_f - K$ poles on the second sheet and K poles on the first sheet. We are not interested in such configurations; we want to put N_f poles at the same point on the same sheet. However, as we discuss now, we actually do not have to worry about the r-branches under the constraint (5.4.3).

The r-branches with different K are different vacua in general. However, under the constraint (5.4.3) these r-branches become all identical because at the branch point $z = a_i^\pm$ there is no distinction between the first and second sheets. This can be easily seen in the matrix model approach. From the equation (5.2.12), the effective glueball superpotential for the two cut model with $N_f - K$ poles at $\widetilde{q}_f$ on the second sheet and K poles at q_f on the first sheet is

$$
\begin{aligned}
W_{\text{eff}} = &-\frac{1}{2}(N_{c,1}\Pi_1 + N_{c,2}\Pi_2) - \frac{1}{2}(N_f - K)\Pi_f^{(2)} - \frac{1}{2}K\Pi_f^{(1)} \\
&+ \frac{1}{2}(2N_c - N_f)W(\Lambda_0) + \frac{1}{2}N_f W(q) \\
&- \pi i(2N_c - N_f)S + 2\pi i\tau_0 S + 2\pi i b_1 S_1,
\end{aligned}
$$

where the periods are defined by

$$
\begin{aligned}
S_i &= \frac{1}{2\pi i}\int_{A_i} R(z)dz, \qquad \Pi_i = 2\int_{a_i^-}^{\Lambda_0} y(z)dz, \\
\Pi_f^{(2)} &= \int_{\widetilde{q}_f}^{\widetilde{\Lambda}_0} y(z)dz, \\
\Pi_f^{(1)} &= \int_{q_f}^{\widetilde{\Lambda}_0} y(z)dz = \left[\int_{q_f}^{\widetilde{q}_f} + \int_{\widetilde{q}_f}^{\widetilde{\Lambda}_0}\right] y(z)dz \equiv \Delta\Pi_f + \Pi_f^{(2)}
\end{aligned}
$$

with $i = 1, 2$. The periods $\Pi_f^{(1)}$, $\Pi_f^{(2)}$ are associated with the poles on the first sheet and the ones on the second sheet, respectively. The contour C_2 for $\Pi_f^{(2)}$ is totally on the second sheet, while the contour C_1 for $\Pi_f^{(1)}$ is from q_f on the first sheet, through a cut, to $\widetilde{\Lambda}_0$ on the second sheet. These contours are shown in Fig. 5.5. This r-branch with K poles on the first sheet can be reached by first starting from the pseudo-confining phase with all N_f poles at $\widetilde{q}_f$ ($K = 0$) and then moving K poles through the cut to q_f. The path along which the poles are moved in this process is the difference in the contours, $C_1 - C_2 \equiv \Delta C$ [15].

When we impose the constraint (5.4.3), then the difference ΔC vanishes (Fig. 5.6). Therefore there is no distinction between C_1, C_2 and hence $\Pi_f^{(1)} = \Pi_f^{(2)}$ for any K. In other words, all K-th branches collapse[16] to the same branch under the constraint (5.4.3).

Now, let us explicitly solve the factorization problem under the constraint (5.4.3), and check that the $S = 0$ solutions exist as advertised before. In solving the factorization problem, we do not have to worry about the r-branches

[15] There is ambiguity in taking ΔC; for example we can take ΔC to go around a_1^+ in Fig. 5.5. However, the difference in $\int_{\Delta C} y(z)dz$ for such different choices of ΔC is $2\pi i n S_1$, $n \in \mathbb{Z}$, which can be absorbed in redefinition of the theta angle and is immaterial.

[16] In fact this collapse was observed in [84, 25] for $SO(N_c)$ and $USp(2N_c)$ gauge groups with massless flavors. We have seen that there are only two branches, i.e., Special branch and Chebyshev branch, which correspond to the baryonic branch and the non-baryonic branch in $U(N_c)$ case.

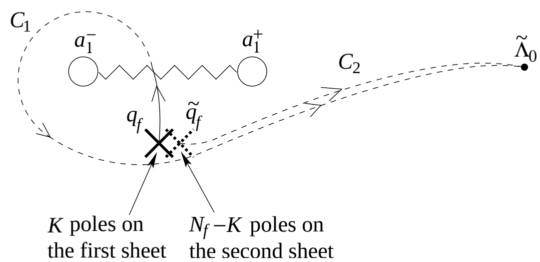


Figure 5.5: *Contours* C_1 *and* C_2 *defining* $\Pi_f^{(1)}$ *and* $\Pi_f^{(2)}$*, respectively. The part of a contour on the first sheet is drawn in a solid line, while the part on the second sheet is drawn in a dashed line. The* $N_f - K$ *poles on the second sheet and the* K *poles on the first sheet are actually on top of each other (more precisely, their projections to the* z*-plane are.)*

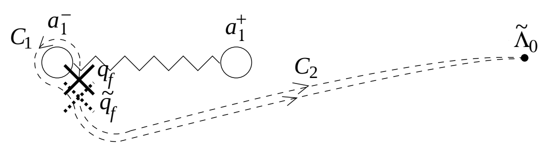


Figure 5.6: *Under the constraint* (5.4.3), *contours* C_1 *and* C_2 *become degenerate:* $C_1 = C_2$.

because there is no distinction among them under the constraint (5.4.3). Then, we compute the exact superpotential using the data from the factorization and reproduce it in matrix model by setting $S = 0$ by hand when $S = 0$ on the gauge theory side.

For simplicity and definiteness, we consider the case of $U(3)$ gauge group henceforth. We consider $N_f < 2N_c$ flavors, namely $1 \leq N_f \leq 5$, because $N_f \geq 2N_c$ cases are not asymptotically free and cannot be treated in the framework of the Seiberg–Witten theory.

5.4.1 Gauge theory computation of superpotential

In this subsection, we solve the factorization equation under the constraint (5.4.3) and compute the exact superpotential, for the system (5.4.1) with $U(3)$ gauge group and with various breaking patterns.

5.4.1.1 Setup

The factorization equation for $U(3)$ theory with N_f flavors with mass $m_f = -z_f$ is given by[17] [81]

$$\begin{aligned} \widetilde{P}_3(\widetilde{z})^2 - 4\Lambda^{6-N_f}(\widetilde{z}-z_f)^{N_f} &= \widetilde{H}_1(\widetilde{z})^2\left[W'(\widetilde{z})^2 + f_1(\widetilde{z})\right] \\ &= \widetilde{H}_1(\widetilde{z})^2(\widetilde{z}-a_1^-)(\widetilde{z}-a_1^+)(\widetilde{z}-a_2^-)(\widetilde{z}-a_2^+), \quad (5.4.4) \end{aligned}$$

where we set $g=1$ for simplicity and $W'(\widetilde{z})$ is given by (5.4.1). The breaking pattern is assumed to be $U(3) \to U(N_{c,1}) \times U(N_{c,2})$ with $N_{c,i} > 0$. Here we used new notations to clarify the shift of the coordinate below. For quantities after the shift, we use letters without tildes. Enforcing the constraint (5.4.3) and shifting $\widetilde{z}$ as $\widetilde{z} = z + a_1^-$, we can rewrite this relation as follows:

$$\begin{aligned} P_3(z)^2 - 4\Lambda^{6-N_f}z^{N_f} &= H_1(z)^2\left[\widetilde{W}'(z)^2 + \widetilde{f}_1(z)\right] \\ &= H_1(z)^2\left[z(z-\widetilde{a}_1^+)(z-\widetilde{a}_2^-)(z-\widetilde{a}_2^+)\right] \\ &\equiv H_1(z)^2\left[z(z^3+Bz^2+Cz+D)\right]. \qquad (5.4.5) \end{aligned}$$

Because of the shift, the polynomials $P_3(z)$ and $H_1(z)$ are different in form from $\widetilde{P}_3(\widetilde{z})$ and $\widetilde{H}_1(\widetilde{z})$ in (5.4.4). We parametrize the polynomials $P_3(z)$ and $H_1(z)$ as

$$P_3(z) = z^3 + az^2 + bz + c, \qquad H_1(z) = z - A. \qquad (5.4.6)$$

[17] From the result of Appendix (5.A.3), the matrix model curve with flavors does not change even for $N_f > N_c$, contrary to [7].

The parameters B, C can be written in terms of the parameters in

$$\widetilde{W}'(z) = W'(\widetilde{z}) = \widetilde{z}(\widetilde{z} + m_A) = (z + a_1^-)(z + a_1^- + m_A) \equiv (z - a_1)(z - a_2)$$

by comparing the coefficients in (5.4.2):

$$\begin{aligned} a_1 &= \frac{-B \mp \sqrt{3\,B^2 - 8\,C}}{4}, \quad a_2 = \frac{-B \pm \sqrt{3\,B^2 - 8\,C}}{4}, \\ \Delta^2 &\equiv (\widetilde{a}_2 - \widetilde{a}_1)^2 = (a_2 - a_1)^2 = \frac{3\,B^2 - 8\,C}{4}. \end{aligned} \tag{5.4.7}$$

The ambiguity in signs in front of the square roots can be fixed by assuming $a_1 < a_2$. Finally we undo the shift by noting that

$$W(\widetilde{z}) = \frac{1}{3}\widetilde{z}^3 + \frac{m_A}{2}\widetilde{z}^2 = \frac{z^3}{3} - (a_1 + a_2)\frac{z^2}{2} + a_1 a_2 z + \frac{1}{6}(a_1^3 - 3a_1^2 a_2). \tag{5.4.8}$$

With all this setup, we can compute the superpotential as follows. First we factorize the curve according to (5.4.5). Then we find the Casimirs U_1, U_2, U_3 from $P_3(z)$ [18] and solve for a_1, a_2 using the last equation of (5.4.7). Finally we put all these quantities into (5.4.8) to get the effective action as

$$W_{\rm low} = \langle \mathrm{Tr}\, W(\Phi) \rangle = U_3 - (a_1 + a_2) U_2 + a_1 a_2 U_1 + \frac{N_c}{6}(a_1^3 - 3a_1^2 a_2). \tag{5.4.9}$$

Here, for a_1, a_2, one can use the first two equations of (5.4.7). We also want to know whether the first cut (the one along the interval $[a_1^-, a_1^+]$) is closed or not; we expect that the cut closes if we try to bring too many poles near the cut. As is obvious from (5.4.5), this can be seen from the value of D. If $D = 0$, the cut is closed, while if $D \neq 0$, the cut is open.

5.4.1.2 The result of factorization problem

We explicitly solved the factorization problem for $U(3)$ gauge theories with $N_f = 1, 2, \ldots, 5$ and summarized the result in Table 1. Let us explain about the table.

[18]From the coefficients of $P_3(z)$, namely a, b and c, one can compute the Casimirs $U_k = \frac{1}{k}\left\langle \mathrm{Tr}\, \Phi^k \right\rangle$ using the quantum modified Newton relation Appendix (5.A.1) as explained in Appendix 5.C.

N_f	breaking pattern $\widehat{U(N_{c,1})} \times U(N_{c,2})$	$N_f \leq N_{c,1}$	$N_{c,1} < N_f < 2N_{c,1}$	$2N_{c,1} \leq N_f$	the first cut is
1	$\widehat{U(2)} \times U(1)$	$\bigcirc$	-	-	open
	$\widehat{U(1)} \times U(2)$	$\bigcirc$	-	-	open
2	$\widehat{U(1)} \times U(2)$	-	-	$\bigcirc$	closed
	$\widehat{U(2)} \times U(1)$	$\bigcirc$	-	-	open
3	$\widehat{U(1)} \times U(2)$	-	-	$\bigcirc$	closed
	$\widehat{U(2)} \times U(1)$	-	$\bigcirc$	-	closed
	$\widehat{U(2)} \times U(1)$	-	$\bigcirc$	-	open
4	$\widehat{U(1)} \times U(2)$	-	-	$\bigcirc$	closed
	$\widehat{U(2)} \times U(1)$	-	-	$\bigcirc$	closed
5	$\widehat{U(1)} \times U(2)$	-	-	$\bigcirc$	closed
	$\widehat{U(2)} \times U(1)$	-	-	$\bigcirc$	closed

Table 5.1: *The result of factorization of curves for $U(3)$ with up to $N_f = 5$ flavors. "$\bigcirc$" denotes which inequality N_f and $N_{c,1}$ satisfy.*

$\widehat{U(N_{c,1})} \times U(N_{c,2})$ denotes the breaking pattern of the $U(3)$ gauge group. The hat on the first factor means that the pole is at one of the branch points of the first cut (Eq. (5.4.3)) which is associated with the first factor $U(N_{c,1})$. Of course this choice is arbitrary and we may as well choose $U(N_{c,2})$, ending up with the same result. Finally, whether the cut is closed or not depends on whether $D = 0$ or not, as explained below (5.4.9).

Now let us look carefully at Table 1, comparing it with the prescription (5.3.24) based on the analysis of the one cut model.

First of all, for $N_f \geq 2N_{c,1}$, the first cut is always closed. We will see below that in these cases with a closed cut the superpotential can be reproduced by setting $S_1 = 0$ by hand in the corresponding matrix model, confirming the prescription (5.3.24) for $N_f > 2N_{c,1}$.

Secondly, the lines for $N_f = 3$ and $\widehat{U(2)} \times U(1)$ correspond to the $N_{c,1} < N_f < 2N_{c,1}$ part of the prescription (5.3.24). There indeed are both an open cut solution and a closed cut solution. We will see below that the superpotential of the closed cut solution can be reproduced by setting $S_1 = 0$ by hand in the corresponding matrix model. On the other hand, the superpotential of the open cut solution can be reproduced by not setting $S_1 = 0$, namely by treating S_1 a dynamical variable and extremizing W_{eff} with respect to it. In fact these two solutions are baryonic branch for a closed cut and non-baryonic branch for an open cut.

Finally, for $N_f \leq N_{c,1}$, the cut is always open, which is also consistent with the prescription (5.3.24). In this case, the superpotential of the open cut solution can be reproduced by extremizing W_{eff} with respect to it, as we will see below. For $N_f = N_{c,1}$ there should be an $S_1 = 0$ ($a_1^- = a_1^+$) solution for some z_f (corresponding to $U(2)$ theory with $N_f = 2$ in the $r = 0$ branch) in the quadratic

case, but under the constraint (5.4.3) we cannot obtain that solution.

Below we present resulting exact superpotentials, for all possible breaking patterns. For simplicity, we do not take care of phase factor of Λ which gives rise to the whole number of vacua. For details of the calculation, see Appendix 5.C.

Results

Definitions: $W_{\rm cl} = -\frac{1}{3}$ for $\widehat{U(1)} \times U(2)$ and $W_{\rm cl} = -\frac{1}{6}$ for $\widehat{U(2)} \times U(1)$. For simplicity we set $g = 1$ and $\Delta = a_2 - a_1 = -m/g = 1$.

- $\boldsymbol{\widehat{U(1)} \times U(2)}$ **with** $\boldsymbol{N_f = 1}$

$$W_{\rm low} = W_{\rm cl} - 2T - \frac{5T^2}{2} + \frac{115T^3}{12} - \frac{245T^4}{4} + \frac{30501T^5}{64} - \frac{12349T^6}{3} + \cdots, \quad T \equiv \Lambda^{\frac{5}{2}}.$$

- $\boldsymbol{\widehat{U(2)} \times U(1)}$ **with** $\boldsymbol{N_f = 1}$

$$W_{\rm low} = W_{\rm cl} - \frac{5T^2}{2} + \frac{5T^3}{3} - \frac{11T^4}{3} + 11T^5 - \frac{235T^6}{6} + \cdots, \quad T \equiv \Lambda^{\frac{5}{3}}.$$

- $\boldsymbol{\widehat{U(1)} \times U(2)}$ **with** $\boldsymbol{N_f = 2}$

$$W_{\rm low} = W_{\rm cl} + 2T^2 - 6T^4 - \frac{32T^6}{3} - 40T^8 - 192T^{10} - \frac{3136T^{12}}{3} + \cdots, \quad T \equiv \Lambda.$$

- $\boldsymbol{\widehat{U(2)} \times U(1)}$ **with** $\boldsymbol{N_f = 2}$

$$W_{\rm low} = W_{\rm cl} - 2T^4 - \frac{16T^6}{3} - 24T^8 - 128T^{10} - \frac{2240T^{12}}{3} + \cdots, \quad T \equiv \Lambda.$$

- $\boldsymbol{\widehat{U(1)} \times U(2)}$ **with** $\boldsymbol{N_f = 3}$

$$W_{\rm low} = W_{\rm cl} + 2T - \frac{19T^2}{2} + \frac{51T^3}{4} + \frac{157T^4}{4} + \frac{5619T^5}{64} + \frac{33T^6}{2} + \cdots, \quad T \equiv \Lambda^{\frac{3}{2}}.$$

- $\widehat{U(2)} \times U(1)$ **with** $N_f = 3$**: two solutions**

$$W_{\text{low,baryonic}} = W_{\text{cl}} + T - \frac{5T^2}{2} - 33T^3 - 543T^4 - 10019T^5 - \frac{396591T^6}{2} + \cdots, \quad T \equiv \Lambda^3.$$

$$W_{\text{low}} = W_{\text{cl}} + \Lambda^3.$$

- $\widehat{U(1)} \times U(2)$ **with** $N_f = 4$

$$W_{\text{low}} = W_{\text{cl}} + 2T - 13T^2 + \frac{176T^3}{3} - 138T^4 + 792T^6 - 9288T^8 + 137376T^{10} - 2286144T^{12} + \cdots, \quad T \equiv \Lambda.$$

- $\widehat{U(2)} \times U(1)$ **with** $N_f = 4$

$$W_{\text{low}} = W_{\text{cl}} + T - 6T^2 - \frac{40T^3}{3} - 56T^4 - 288T^5 - \frac{4928T^6}{3} - 9984T^7 - 63360T^8 - \frac{1244672T^9}{3} - 2782208T^{10} - 19009536T^{11} \cdots, \quad T \equiv \Lambda^2.$$

- $\widehat{U(1)} \times U(2)$ **with** $N_f = 5$

$$W_{\text{low}} = W_{\text{cl}} - 2T - \frac{33T^2}{2} - \frac{1525T^3}{12} - \frac{3387T^4}{4} - \frac{314955T^5}{64} - \frac{74767T^6}{3} + \cdots, \quad T \equiv \Lambda^{\frac{1}{2}}.$$

- $\widehat{U(2)} \times U(1)$ **with** $N_f = 5$

$$W_{\text{low}} = W_{\text{cl}} + T - \frac{19T^2}{2} + \frac{154T^3}{3} - 132T^4 + 828T^6 + \cdots, \quad T \equiv \Lambda.$$

5.4.2 Matrix model computation of superpotential

In this subsection we compute the superpotential of the system (5.4.1) in the framework of [35]. If all the N_f flavors have the same mass $m_f = -z_f$, the effective

glueball superpotential $W_{\rm eff}(S_j)$ for the pseudo-confining phase with breaking pattern $U(N_c) \to \prod_{i=1}^n U(N_{c,i})$, $\sum_{i=1}^n N_{c,i} = N_c$ is, from (5.2.12),

$$\begin{aligned} W_{\rm eff}(S_j) = & -\frac{1}{2}\sum_{i=1}^{n} N_{c,i}\Pi_i - \frac{N_f}{2}\Pi_f + \left(N_c - \frac{N_f}{2}\right) W(\Lambda_0) + \frac{N_f}{2} W(z_f) \\ & - 2\pi i \left(N_c - \frac{N_f}{2}\right) S + 2\pi i \tau_0 S + 2\pi i \sum_{i=1}^{n-1} b_i S_i, \end{aligned} \tag{5.4.10}$$

where the periods associated with adjoint and fundamentals are defined by

$$\Pi_i(S_j) \equiv 2\int_{a_i^-}^{\Lambda_0} y(z)dz, \qquad \Pi_f(S_j) \equiv \int_{\tilde{z}_f}^{\tilde{\Lambda}_0} y(z)dz = -\int_{z_f}^{\Lambda_0} y(z)dz,$$
$$y(z) = \sqrt{W'(z)^2 + f_1(z)}\ .$$

For cubic tree level superpotential (5.4.1), the periods $\Pi_{1,2}(S_j)$ were computed by explicitly evaluating the period integrals by power expansion in [10], as

$$\begin{aligned} \frac{\Pi_1}{2g\Delta^3} = & \frac{1}{g\Delta^3}[W(\Lambda_0) - W(a_1)] + s_1\left[1 + \log\left(\frac{\lambda_0^2}{s_1}\right)\right] + 2s_2 \log\lambda_0 \\ & + (-2s_1^2 + 10 s_1 s_2 - 5 s_2^2) + \left(-\frac{32}{3}s_1^3 + 91 s_1^2 s_2 - 118 s_1 s_2^2 + \frac{91}{3}s_2^3\right) \\ & + \left(-\frac{280}{3}s_1^4 + \frac{3484}{3}s_1^3 s_2 - 2636 s_1^2 s_2^2 + \frac{5272}{3}s_1 s_2^3 - \frac{871}{3}s_2^4\right) + \cdots, \\ \frac{\Pi_2}{2g\Delta^3} = & \frac{1}{g\Delta^3}[W(\Lambda_0) - W(a_2)] + s_2\left[1 + \log\left(\frac{\lambda_0^2}{-s_2}\right)\right] + 2s_1 \log\lambda_0 \\ & + (2s_2^2 - 10 s_1 s_2 + 5 s_1^2) + \left(-\frac{32}{3}s_2^3 + 91 s_1 s_2^2 - 118 s_1^2 s_2 + \frac{91}{3}s_1^3\right) \\ & + \left(\frac{280}{3}s_2^4 - \frac{3484}{3}s_1 s_2^3 + 2636 s_1^2 s_2^2 - \frac{5272}{3}s_1^3 s_2 + \frac{871}{3}s_1^4\right) + \cdots, \end{aligned} \tag{5.4.11}$$

where $\Delta \equiv a_2 - a_1$, $s_i \equiv S_i/g\Delta^3$, and $\lambda_0 \equiv \Lambda_0/\Delta$.

Under the constraint (5.4.3), the contours defining Π_1 and Π_f coincide, so

$$\frac{1}{2}\Pi_1 = -\Pi_f = \int_{z_f = a_1^-}^{\Lambda_0} y(z)dz. \tag{5.4.12}$$

Using this, we can rewrite (5.4.10) as

$$\begin{aligned} W_{\text{eff}}(S_1, S_2) = &\Big[N_{c,1}W(a_1) + N_{c,2}W(a_2)\Big] - \frac{N_f}{2}\Big[W(a_1) - W(z_f)\Big] \\ &- \widetilde{N}_{c,1}\Big[\frac{1}{2}\Pi_1 - W(\Lambda_0) + W(a_1)\Big] - N_{c,2}\Big[\frac{1}{2}\Pi_2 - W(\Lambda_0) + W(a_2)\Big] \\ &- 2\pi i(\widetilde{N}_{c,1} + N_{c,2})S + 2\pi i\tau_0 S + 2\pi i b_1 S_1. \end{aligned} \tag{5.4.13}$$

Here we rearranged the terms taking into account the fact that the periods take the form $\frac{1}{2}\Pi_i = W(\Lambda_0) - W(a_i) +$ (quantum correction of order $\mathcal{O}(S_i)$), and also the fact that we are considering $z_f = a_1^- \simeq a_1$ (thus the second term). The first line corresponds to the classical contribution, while the second and third lines correspond to quantum correction. Furthermore, we defined $\widetilde{N}_{c,1} \equiv N_{c,1} - N_f/2$.

We would like to extremize this W_{eff} (5.4.13) with respect to $S_{1,2}$, and compute the low energy superpotential that can be compared with the W_{low} obtained in the previous subsection using gauge theory methods. In doing that, one should be careful to the fact that one should treat the mass z_f as an external parameter which is independent of $S_{1,2}$ although we are imposing the constraint (5.4.3), $z_f = a_1^- = a_1^-(S_1, S_2)$. Where is the z_f dependence in (5.4.13)? Firstly, z_f appears explicitly in the second term in (5.4.13). Therefore, when we differentiate W_{eff} with respect to $S_{1,2}$, we should exclude this term. Secondly, there is a more implicit dependence on z_f in $\Pi_f = -\int_{z_f}^{\Lambda_0} y(z)dz$, which we replaced with $-\Pi_1/2$ using (5.4.12). If we forget to treat z_f as independent of S_i, then we get an apparently unwanted, extra contribution as $\frac{\partial \Pi_f}{\partial S_i} = -\int_{z_f}^{\Lambda_0} \frac{\partial y(z)}{\partial S_i}dz - \frac{\partial z_f}{\partial S_i} \cdot y(z)|_{z=z_f}$. However, this last term actually does not make difference because

$$y(z)|_{z=z_f} = g\,\sqrt{(z-z_f)(z-a_1^+)(z-a_2^-)(z-a_2^+)}\,\Big|_{z=z_f} = 0.$$

Therefore what one should do is: i) plug the expression (5.4.11) into (5.4.13), ii) solve the equation of motion for $S_{1,2}$ using (5.4.13) without the second term, and

then iii) substitute back the value of $S_{1,2}$ into (5.4.13), now with the second term included.

Solving the equation of motion can be done by first writing the Veneziano–Yankielowicz term (log and linear terms) as

$$\begin{aligned} -\widetilde{N}_{c,1}\Big[\frac{1}{2}\Pi_1 - W(\Lambda_0) + W(a_1)\Big] - N_{c,2}\Big[\frac{1}{2}\Pi_2 - W(\Lambda_0) + W(a_2)\Big] \\ - 2\pi i(\widetilde{N}_{c,1} + N_{c,2})S + 2\pi i\tau_0 S + 2\pi i b_1 S_1 \\ = g\Delta^3\Big\{\widetilde{N}_{c,1}s_1[1 - \log(s_1/\lambda_1^3)] + N_{c,2}s_2[1 - \log(s_2/\lambda_2^3)] + \mathcal{O}(s_i^2)\Big\}, \end{aligned}$$

where

$$\begin{aligned} \lambda_1^{3\widetilde{N}_{c,1}} &= \lambda_0^{2(\widetilde{N}_{c,1}+2N_{c,2})} e^{2\pi i(\widetilde{N}_{c,1}+N_{c,2})-2\pi i\tau_0-2\pi i b_1}, \\ \lambda_2^{3N_{c,2}} &= (-1)^{N_{c,2}}\lambda_0^{2\widetilde{N}_{c,1}+2N_{c,2}} e^{2\pi i(\widetilde{N}_{c,1}+N_{c,2})-2\pi i\tau_0}, \end{aligned}$$

and then solving the equation of motion perturbatively in $\lambda_{1,2}$. In this way, one can straightforwardly reproduce the results obtained in the previous section in the case with the first cut *open*. In the case with the first cut *closed*, in order to reproduce the results in the previous section, one should first set $S_1 = 0$ by hand, and then extremize W_{eff} with respect to the remaining dynamical variable S_2.

Following the procedure above, we checked explicitly that extremizing $W_{\text{eff}}(S_1, S_2)$ (open cut) or $W_{\text{eff}}(S_1{=}0, S_2)$ (closed cut) reproduces the W_{low} up the order presented in the previous section, for all breaking patterns for $U(3)$ theory.

In the above, we concentrated the explicit calculations of effective superpotentials in $U(3)$ theory with cubic tree level superpotential. These explicit examples are useful to see that the prescription (5.3.24) really works; one can first determine using factorization method when we should set $S_i = 0$ by hand, and then explicitly check that the superpotential obtained by gauge theory can be reproduced by matrix model.

However, if one wants only to show the equality of the two effective superpotentials on the gauge theory and matrix model sides, one can actually prove it in general cases. In Appendix 5.B, we prove this equivalence for $U(N_c)$ gauge theory with an degree $(k+1)$ tree level superpotential where $k+1 < N_c$. There, we show the following: if there are solutions to the factorization problem with some cuts closed, then the superpotential $W_{\rm low}$ of the gauge theory can be reproduced by extremizing the glueball superpotential $W_{\rm eff}(S_i)$ on the matrix model side, after setting the corresponding glueball fields S_i to zero by hand. Note that, on the matrix side we do not know when we should set S_i to zero *a priori*; we can always set S_i to zero in matrix model, but that does not necessarily correspond to a physical solution on the gauge theory side that solves the factorization constraint.

5.5 Conclusion and some remarks

In this paper, taking $\mathcal{N}=1$ $U(N_c)$ gauge theory with an adjoint and flavors, we studied the on-shell process of passing N_f flavor poles on top of each other on the second sheet through a cut onto the first sheet. This corresponds to a continuous transition from the pseudo-confining phase with $U(N_c)$ unbroken to the Higgs phase with $U(N_c-N_f)$ unbroken (we are focusing on one cut). We confirmed the conjecture of [35] that for $N_f < N_c$ the poles can go all the way to infinity on the first sheet, while for $N_f \geq N_c$ there is obstruction. There are two types of obstructions: the first one is that the cut rotates, catches poles and sends them back to the first sheet, while the second one is that the cut closes up before poles reach it. The first obstruction occurs for $N_c \leq N_f < 2N_c$ whereas the second one occurs for $N_c < N_f$.

If a cut closes up, the corresponding glueball S vanishes, which means that the

$U(N_c)$ group is completely broken down. This can happen only by condensation of a charged massless degree of freedom, which is missing in the matrix model description of the system. With a massless degree of freedom missing in the description, the $S = 0$ solution should be singular in matrix model in some sense. Indeed, we found that the $S = 0$ solution of the gauge theory does not satisfy the equation of motion in matrix model (with an exception of the $N_f = N_c$ case, where the $S = 0$ solution does satisfy the equation of motion). How to cure this defect of matrix model is simple — the only thing the missing massless degree of freedom does is to make $S = 0$ a solution, so we just set $S = 0$ by hand in matrix model. We gave a precise prescription (5.3.24) when we should do this, i.e., *in the baryonic branch for* $N_{c,i} \le N_f < 2N_{c,i}$ *and in the* $r = N_{c,i}$ *non-baryonic branch*, and checked it with specific examples.

The string theory origin of the massless degree of freedom can be conjectured by generalizing the argument in [87]. We argued that it should be the D3-brane wrapping the blown up S^3, along with fundamental strings emanating from it and ending on the noncompact D5-branes in the Calabi–Yau geometry.

Although we checked that the prescription works, the string theory picture of the $S = 0$ solution needs further refinement, which we leave for future research. For example, although we argued that some extra degree of freedom makes $S = 0$ a solution, we do not have the precise form of the superpotential including that extra field. It is desirable to derive it and show that $S = 0$ is indeed a solution, as was done in [87] in the case without flavors. Furthermore, we saw that there is an on-shell $S = 0$ solution for $N_f = N_c$. Although this solution solves the equation of motion in matrix model, there should be a massless field behind the scene. It is interesting to look for the nature of this degree of freedom. It cannot be the D3-branes with fundamental strings emanating from it, since for this solution the

noncompact D5-branes are at finite distance from the collapsed S^3 and the 3-5 strings are massive. Finally, we found that the $S = 0$ solution is in the baryonic branch. It would be interesting to ask if one can describe the baryonic branch in the matrix model framework by adding some extra degrees of freedom.

In the following, we study some aspects of the theory, which we could not discuss so far. We will discuss generalization to $SO(N_c)$ and $USp(2N_c)$ gauge groups by computing the effective superpotentials with quadratic tree level superpotential.

5.5.1 $SO(N_c)$ theory with flavors

Here we consider the one cut model for $SO(N_c)$ gauge theory with N_f flavors. The tree level superpotential of the theory is obtained from $\mathcal{N} = 2$ SQCD by adding the mass m_A for the adjoint scalar Φ

$$W_{\rm tree} = \frac{m_A}{2} \mathrm{Tr}\, \Phi^2 + Q^f \Phi Q^{f'} J_{ff'} + Q^f \widetilde{m}_{ff'} Q^{f'}. \qquad (5.5.1)$$

where $f = 1, 2, \cdots, 2N_f$ and the symplectic metric $J_{ff'}$ and mass matrix for quark $\widetilde{m}_{ff'}$ are given by

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes \mathbb{1}_{N_f \times N_f}, \qquad \widetilde{m} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \mathrm{diag}(m_1, \cdots, m_{N_f}) \; .$$

For this simple case the matrix model curve is given by

$$y(z)^2 = m_A^2 \left(z^2 - 4\mu^2 \right). \qquad (5.5.2)$$

This Riemann surface is a double cover of the complex z-plane branched at the roots of y_m^2 (that is $z = \pm 2\mu$).

The effective superpotential receives contributions from both the sphere and the disk amplitudes in the matrix model [47] and the explicit form was given in

[35] for $U(N_c)$ gauge theory with flavors. Now we apply this procedure to our $SO(N_c)$ gauge theory with flavors and it turns out the following expression

$$\begin{aligned} W_{\text{eff}} = -\frac{1}{2}\left(\sum_{i=-n}^{n} N_{c,i} - 2\right)\int_{\widehat{B}_i^r} y(z)dz - \frac{1}{4}\sum_{I=1}^{2N_f}\int_{\widetilde{q}_I}^{\widetilde{\Lambda}_0} y(z)dz \\ + \frac{1}{2}\left(2N_c - 4 - 2N_f\right)W(\Lambda_0) + \frac{1}{2}\sum_{I=1}^{2N_f} W(z_I) \\ - \pi i\left(2N_c - 4 - 2N_f\right)S + 2\pi i\tau_0 S + 2\pi i\sum_{i=1}^{n} b_i S_i \end{aligned}$$

where $S = S_0 + 2\sum_{i=1}^{n} S_i$ and z_I is the root of

$$B(z) = \det m(z) = \prod_{I=1}^{N_f}\left(z^2 - z_I^2\right) \ .$$

Since the curve (5.5.2) is same as the one (5.3.1) of $U(N_c)$ gauge theory, we can use the integral results given there to write down the effective superpotential as

$$\begin{aligned} W_{\text{eff}} = S\left[\frac{(N_c-2)}{2} + \log\left(\frac{2^{\frac{N_c-2}{2}} m_A^{\frac{N_c-2}{2}} \Lambda^{N_c-2-N_f}\det z}{S^{\frac{N_c-2}{2}}}\right)\right] \\ - S\sum_{I=1, r_I=0}^{N_f}\left[-\log\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{2S}{m_A z_I^2}}\right) + \frac{m_A z_I^2}{2S}\left(\sqrt{1 - \frac{2S}{m_A z_I^2}} - 1\right) + \frac{1}{2}\right] \\ - S\sum_{I=1, r_I=1}^{N_f}\left[-\log\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{2S}{m_A z_I^2}}\right) + \frac{m_A z_I^2}{2S}\left(-\sqrt{1 - \frac{2S}{m_A z_I^2}} - 1\right) + \frac{1}{2}\right]. \end{aligned}$$

We can solve $M(z)$ and $T(z)$ as did for $U(N_c)$ gauge theory. For simplicity we take all $r_I = 0$, i.e., all poles at the second sheet. For the I-th block diagonal matrix element of $M(z)$ ($I = 1, \cdots, N_f$) it is given by

$$M_I(z) = \begin{pmatrix} 0 & -\frac{R(z) - R(q_I = m_I)}{z - m_I} \\ \frac{R(z) - R(q_I = -m_I)}{z + m_I} & 0 \end{pmatrix}$$

where

$$R(z) = m_A \left(z - \sqrt{z^2 - 4\mu^2} \right).$$

Expanding $M_I(z)$ in the series of z we can find $\left\langle Q^f \Phi Q^{f'} J_{ff'} + Q^f \widetilde{m}_{ff'} Q^{f'} \right\rangle = 2N_f S$.

The gauge invariant operator $T(z)$ can be constructed similarly as follows:

$$T(z) = \frac{B'(z)}{2B(z)} - \sum_{I=1}^{N_f} \frac{y(q_I) z_I}{y(z)\left(z^2 - z_I^2\right)} + \frac{c(z)}{y(z)} - \frac{2}{z}\frac{R(z)}{y(z)} \tag{5.5.3}$$

where

$$c(z) = \left\langle \mathrm{Tr} \frac{W'(z) - W'(\Phi)}{z - \Phi} \right\rangle - \sum_{I=1}^{N_f} \frac{z W'(z) - z_I W'(z_I)}{\left(z^2 - z_I^2\right)}.$$

For the theory without the quarks, the Konishi anomaly was derived in [30, 31, 97]. The last term in (5.5.3) reflects the action of orientifold. For our example we have

$$c(z) = m_A \left(N_c - N_f \right).$$

and

$$\begin{aligned} T(z) = {} & \frac{1}{z} N_c + \frac{1}{z^3} \left[\sum_{I=1}^{N_f} z_I \left(z_I - \sqrt{z_I^2 - 4\mu^2} \right) + 2\mu^2 \left(N_c - 2 - N_f \right) \right] \\ & + \frac{1}{z^5} \left[\sum_{I=1}^{N_f} z_I^4 - \sum_{I=1}^{N_f} \sqrt{z_I^2 - 4\mu^2}\, z_I \left(z_I^2 + 2\mu^2 \right) + 6\mu^4 \left(N_c - N_f \right) - 12\mu^4 \right] \\ & + \mathcal{O}\left(1/z^7 \right), \end{aligned}$$

where for equal mass of flavor, we get

$$\begin{aligned} \left\langle \mathrm{Tr}\, \Phi^2 \right\rangle &= N_f q \left(q - \sqrt{q^2 - 4\mu^2} \right) + 2\mu^2 \left(N_c - 2 - N_f \right) \\ \left\langle \mathrm{Tr}\, \Phi^4 \right\rangle &= N_f q^4 - N_f \sqrt{q^2 - 4\mu^2}\, q \left(q^2 + 2\mu^2 \right) + 6\mu^4 \left(N_c - 2 - N_f \right). \end{aligned}$$

Let us assume the mass of flavors are the same and K of them (in this case, $r_I = 1$) locate at the first sheet while the remainder $(N_f - K)$ where $r_I = 0$ are at the second sheet. Then from the effective superpotential, it is ready to extremize this with respect to the glueball field S [19]

$$
\begin{aligned}
0 = \log\left(\frac{\Lambda_1^{\frac{3}{2}(N_c-2)-N_f}}{S^{\frac{N_c-2}{2}}}\right) \\
+ K \log\left(\frac{z_f - \sqrt{{z_f}^2 - \frac{2S}{m_A}}}{2}\right) + (N_f - K)\log\left(\frac{z_f + \sqrt{{z_f}^2 - \frac{2S}{m_A}}}{2}\right). \quad (5.5.4)
\end{aligned}
$$

Or by rescaling the fields $\widehat{S} = \frac{S}{2m_A\Lambda^2}$, $\widehat{z}_f = \frac{z_f}{\Lambda}$ one gets the solution and consider for $K = 0$ case

$$
1 = \widehat{S}^{-\frac{N_c-2}{2}}\left(\frac{\widehat{z}_f + \sqrt{\widehat{z}_f^2 - 4\widehat{S}}}{2}\right)^{N_f}.
$$

This equation is the same as the one in $U(N_c)$ case with $N_c \to \frac{N_c-2}{2}$, so the discussion of passing poles will go through without modification and the result is when $N_f \geq \frac{N_c-2}{2}$, the on-shell poles at the second plane cannot pass the cut to reach the first sheet far away from the cut.

By using the condition (5.5.4), one gets the on-shell effective superpotential

$$
\begin{aligned}
W_{\text{eff,on-shell}} = \frac{1}{2}(N_c - 2 - N_f)\, S \\
+ \frac{1}{2} m_A {z_f}^2 \left(N_f + (2K - N_f)\sqrt{1 - \frac{2S}{m_A {z_f}^2}}\right).
\end{aligned}
$$

It can be checked that this is the same as $\frac{1}{2} m_A \langle \mathrm{Tr}\, \Phi^2 \rangle$.

[19] One can easily check that this equation with parameters (N_c, N_f, K) is equivalent to the one with parameters $(N_c - 2r, N_f - 2r, K - r)$. The equation of motion for glueball field is the same. Since the equation of motion for both r-th Higgs branch and $(N_f - r)$-th Higgs branch is the same, one expects that both branches have some relation. By redefinition of $S \to \frac{4m_A^2\Lambda^4}{S} \equiv \widetilde{S}, z_f \to \frac{2m_A\Lambda^2}{S} z_f \equiv \widetilde{z_f}$ we get the final relation between K Higgs branch and $(N_f - K)$ Higgs branch.

5.5.2 $USp(2N_c)$ theory with flavors

For the $USp(2N_c)$ gauge theory with N_f flavors we will sketch the discussion because most of them is similar to $SO(N_c)$ gauge theory. The tree level superpotential is given by (5.5.1) but with J the symplectic metric, and $\widetilde{m}_{ff'}$ the quark mass given by

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \otimes \mathbb{1}_{N_c \times N_c}, \qquad \widetilde{m} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \otimes \mathrm{diag}(m_1, \cdots, m_{N_f}) \; .$$

We parametrize the matrix model curve and the resolvent $R(z)$ as

$$\begin{aligned} y_m^2 &= W'(z)^2 + f(z) = m_A^2 \left(z^2 + 4\mu^2\right), \\ R(z) &= m_A \left(z - \sqrt{z^2 + 4\mu^2}\right). \end{aligned}$$

The effective superpotential is given by

$$\begin{aligned} W_{\text{eff}} &= S\,(N_c+1)\left[1 + \log\left(\frac{\tilde{\Lambda}^3}{S}\right)\right] \\ &- S \sum_{I=1, r_I=0}^{N_f} \left[-\log\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 + \frac{2S}{m_A z_I^2}}\right) - \frac{m_A z_I^2}{2S}\left(\sqrt{1 + \frac{2S}{m_A z_I^2}} - 1\right) + \frac{1}{2}\right] \\ &- S \sum_{I=1, r_I=1}^{N_f} \left[-\log\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 + \frac{2S}{m_A z_I^2}}\right) - \frac{m_A z_I^2}{2S}\left(-\sqrt{1 + \frac{2S}{m_A z_I^2}} - 1\right) + \frac{1}{2}\right] \end{aligned}$$

and $T(z)$ is given by

$$T(z) = \frac{B'(z)}{2B(z)} - \sum_{I=1}^{N_f} \frac{y(q_I) z_I}{y(z)\,(z^2 - z_I^2)} + \frac{c(z)}{y(z)} + \frac{2}{z}\frac{R(z)}{y(z)}.$$

Note the last term (different sign) compared with the $SO(N_c)$ gauge theory. From the solution of $M(z)$, we can show that although

$$\left\langle Q^f (\Phi J)_{ff'} Q^{f'} + Q^f \widetilde{m}_{ff'} Q^{f'} J \right\rangle = 2N_f S \neq 0,$$

we still have on-shell relation $\frac{1}{2} m_A \left\langle \mathrm{Tr}\, \Phi^2 \right\rangle = W_{\text{eff}}$.

The equation of motion is given by [20]

$$0 = \log \widehat{S}^{-N_c-1} + K \log\left(\frac{\widehat{z}_f - \sqrt{\widehat{z}_f^2 + 4\widehat{S}}}{2}\right) + (N_f - K)\log\left(\frac{\widehat{z}_f + \sqrt{\widehat{z}_f^2 + 4\widehat{S}}}{2}\right)$$

where $\widehat{S} = \frac{S}{2m_A\Lambda^2}$, $\widehat{z}_f = \frac{z_f}{\Lambda}$. From this we can read out the following result: when $N_f \geq N_c + 1$, the on-shell poles at the second sheet cannot pass through the cut to reach the first sheet far away from the cut.

Appendix

5.A On matrix model curve with $N_f(> N_c)$ flavors

In this Appendix we prove by strong coupling analysis that matrix model curve corresponding to $U(N_c)$ supersymmetric gauge theory with N_f flavors is exactly the same as the one without flavors when the degree $(k+1)$ of tree level superpotential W_{tree} is less than N_c [21]. This was first proved in [80] but the derivation was valid only for the range $N_f < N_c$. Then in [81], the proof was extended to the cases with the range $2N_c > N_f \geq N_c$. However, in [81], the characteristic function $P_{N_c}(x)$ was defined by $P_{N_c}(x) = \det(x - \Phi)$, without taking into account the possible quantum corrections due to flavors. In consequence, it appeared that

[20] One can easily check that this equation with parameters $(2N_c, N_f, K)$ is equivalent to the one with parameters $(2N_c - 2r, N_f - 2r, K - r)$. In other words, the equation of motion for glueball field is the same. Since the equation of motion for both r-th Higgs branch and $(K-r)$-th Higgs branch is equivalent to each other, one expects that both branches have some relation.

[21] The generalized Konishi anomaly equation of $R(z)$ given in (5.2.6) is same with or without flavors, so the form of the solution is the same for gauge theory with or without flavor. In this Appendix we use another method to prove this result.

the matrix model curve is changed by addition of flavors. In this Appendix, we use the definition of $P_{N_c}(x)$ proposed in Eq. (C.2) of [35]:

$$P_{N_c}(x) = x^{N_c}\exp\left(-\sum_{i=1}^{\infty}\frac{U_i}{x^i}\right) + \Lambda^{2N_c-N_f}\frac{\widehat{B}(x)}{x^{N_c}}\exp\left(\sum_{i=1}^{\infty}\frac{U_i}{x^i}\right), \qquad (5.A.1)$$

which incorporates quantum corrections and reduces to $P_{N_c}(x) = \det(x-\Phi)$ for $N_f = 0$, and see clearly that the matrix model curve is not changed, even when the number of flavors is more than N_c. Since $P_{N_c}(x)$ is a polynomial in x, (5.A.1) can be used to express U_r with $r > N_c$ in terms of U_r with $r \leq N_c$ by imposing the vanishing of the negative power terms in x.

Assuming that the unbroken gauge group at low energy is $U(1)^n$ with $n \leq k$, the factorization form of Seiberg–Witten curve can be written as,

$$P_{N_c}^2(x) - 4\Lambda^{2N_c-N_f}\widehat{B}(x) = H_{N_c-n}^2(x)F_{2n}(x).$$

The effective superpotential with this double root constraint can be written as follows [22].

$$W_{\text{eff}} = \sum_{r=0}^{k} g_r U_{r+1} + \sum_{i=1}^{N_c-n}\left[L_i \oint \frac{P_{N_c}(x) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(x)}}{x-p_i}dx \right.$$
$$\left. + B_i \oint \frac{P_{N_c}(x) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(x)}}{(x-p_i)^2}dx\right].$$

The equations of motion for B_i and p_i are given as follows respectively:

$$0 = \oint \frac{P_{N_c}(x) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(x)}}{(x-p_i)^2}dx,$$
$$0 = 2B_i \oint \frac{P_{N_c}(x) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(x)}}{(x-p_i)^3}dx.$$

[22]If we want to generalize this proof to more general cases in which $k+1$ is greater than and equals to N_c, we have to take care more constraints like Appendix A in [79], which should be straightforward.

Assuming that the factorization form does not have any triple or higher roots, we obtain $B_i = 0$ at the level of equation of motion. Next we consider the equation of motion for U_r:

$$0 = g_{r-1} + \sum_{i=1}^{N_c-n} \oint \left[\frac{P_{N_c}}{x^r} - 2\frac{x^{N_c}}{x^r} \exp\left(-\sum_{i=1}^{\infty} \frac{U_i}{x^i} \right) \right] \frac{L_i}{x - p_i} dx$$

where we used $B_i = 0$ and (5.A.1) to evaluate $\frac{\partial P_{N_c}}{\partial U_r}$. Now, as in [79], we multiply this by z^{r-1} and sum over r.

$$W'(z) = -\oint \frac{P_{N_c}}{x-z} \sum_{i=1}^{N_c-n} \frac{L_i}{x-p_i} dx + \oint \frac{2x^{N_c}}{x-z} \exp\left(-\sum_{k=1}^{\infty} \frac{U_k}{x^k} \right) \sum_{i=1}^{N_c-n} \frac{L_i}{x-p_i} dx.$$

Defining the polynomial $Q(x)$ in terms of

$$\sum_{i=1}^{N_c-n} \frac{L_i}{x-p_i} = \frac{Q(x)}{H_{N_c-n}(x)}, \tag{5.A.2}$$

and also using (5.A.1) and factorization form, we obtain

$$\begin{aligned} W'(z) &= -\oint \frac{P_{N_c}}{x-z} \frac{Q(x)}{H_{N_c-n}(x)} dx + \oint \frac{P_{N_c}}{x-z} \frac{Q(x)}{H_{N_c-n}(x)} dx + \oint \frac{Q(x)\sqrt{F_{2n}(x)}}{x-z} dx \\ &= \oint \frac{Q(x)\sqrt{F_{2n}(x)}}{x-z} dx. \end{aligned}$$

This is nothing but (2.37) in [79]. Since $W'(z)$ is a polynomial of degree k, the Q should be a polynomial of degree $(k-n)$. Therefore, we conclude that the matrix model curve is not changed by addition of flavors:

$$y_m^2 = F_{2n}(x) Q_{k-n}^2(x) = W_k'(x)^2 + \mathcal{O}(x^{k-1}). \tag{5.A.3}$$

5.B Equivalence between W_{low} and $W_{\text{eff}}(\langle S_i \rangle)$ with flavors

In this Appendix we prove the equivalence W_{low} in $U(N_c)$ gauge theory with $W_{\text{eff}}(\langle S_i \rangle)$ in corresponding dual geometry when some of the branch cuts on the

Riemann surface are closed and the degree $(k+1)$ of the tree level superpotential W_{tree} is less than N_c. This was first proved in [80], however, the proof was only applicable in the $N_f < N_c$ cases. Especially, the field theory analysis in [80] did not work for $N_c \leq N_f < 2N_c$ cases. Furthermore, as we saw in the main text, for some particular choices of z_I (position of the flavor poles), extra double roots appear in the factorization problem. In section 5.4, we dealt with $U(3)$ with cubic tree level superpotential and saw the equivalence of two effective superpotentials for such special situations. To include these cases we are interested in the Riemann surface that has some closed branch cuts. Therefore our proof is applicable for $U(N_c)$ gauge theories with W_{tree} of degree $k+1$ $(< N_c)$ in which some of branch cuts are closed and number of flavors is in the range $N_c \leq N_f < 2N_c$. In addition, we restrict our discussion to the Coulomb phase.

In the discussion below, we follow the strategy developed by Cachazo and Vafa in [76] and use (5.A.1) as the definition of $P_{N_c}(x)$. We have only to show the two relations:

$$W_{\text{low}}(g_r, z_I, \Lambda)\big|_{\Lambda\to 0} = W_{\text{eff}}(\langle S_i \rangle)\big|_{\Lambda\to 0}, \tag{5.B.1}$$

$$\frac{\partial W_{\text{low}}(g_r, z_I, \Lambda)}{\partial \Lambda} = \frac{\partial W_{\text{eff}}(\langle S_i \rangle)}{\partial \Lambda}, \tag{5.B.2}$$

the equivalence of two effective superpotentials in the classical limit and that of the derivatives of the superpotentials with respect to Λ.

5.B.1 Field theory analysis

Let k be the order of W'_{tree} and $n(\leq k)$ be the number of $U(1)$ at low energy. Since we are interested in cases with degenerate branch cuts, let us consider the

following factorization form [23]:

$$P_{N_c}(x)^2 - 4\Lambda^{2N_c-N_f}\widehat{B}(x) = F_{2n}(x)\left[Q_{k-n}(x)\widetilde{H}_{N_c-k}(x)\right]^2$$
$$\equiv F_{2n}(x)\left[H_{N_c-n}(x)\right]^2, \tag{5.B.3}$$
$$W'(x)^2 + f_{k-1}(x) = F_{2n}(x)Q_{k-n}(x)^2.$$

If k equals to n, all the branch cuts in $F_{2k}(x)$ are open. The low energy effective superpotential is given by

$$W_{\text{low}} = \sum_{r=1}^{n+1} g_r U_r + \sum_{i=1}^{l}\left[L_i\left(P_{N_c}(p_i) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(p_i)}\right)\right.$$
$$\left. + Q_i\frac{\partial}{\partial p_i}\left(P_{N_c}(p_i) - 2\epsilon_i\Lambda^{N_c-\frac{N_f}{2}}\sqrt{\widehat{B}(p_i)}\right)\right],$$

where $l \equiv N - n$ and $P_{N_c}(x)$ is defined by (C.3) or (C.4) in [35],

$$P_{N_c}(x) = \langle\det(x-\Phi)\rangle + \left[\Lambda^{2N_c-N_f}\frac{\widehat{B}(x)}{x^{N_c}}\exp\left(\sum_{i=1}^{\infty}\frac{U_i}{x^i}\right)\right]_+. \tag{5.B.4}$$

The second term is specific to the $N_f \geq N_c$ case, representing quantum correction. Define

$$K(x) \equiv \left[\Lambda^{2N_c-N_f}\frac{\widehat{B}(x)}{x^{N_c}}\exp\left(\sum_{i=1}^{\infty}\frac{U_i}{x^i}\right)\right]_+.$$

The first term in (5.B.4) can be represented as $\langle\det(x-\Phi)\rangle \equiv \sum_{k=0}^{N_c} x^{N_c-k}s_k$. The relation between U_i's and s_k's are given by the ordinary Newton relation, $ks_k + \sum_{r=1}^{k} rU_r s_{k-s} = 0$. From the variations of W_{low} with respect to p_i and Q_i, we conclude that $Q_i = 0$ at the level of the equation of motion. In addition, the variation of W_{low} with respect to U_r leads to

$$g_r = -\sum_{i=1}^{l} L_i\frac{\partial P_{N_c}(p_i)}{\partial U_r} = \sum_{i=1}^{l}\sum_{j=0}^{N_c} L_i p_i^{N_c-j}s_{j-r} - \sum_{i=1}^{l} L_i\frac{\partial K(p_i)}{\partial U_r}$$
$$= \sum_{i=1}^{l}\sum_{j=0}^{N_c} L_i p_i^{N_c-j}s_{j-r} - \sum_{i=1}^{l} L_i\left[\Lambda^{2N_c-N_f}\frac{\widehat{B}(p_i)}{p_i^{N_c+r}}\exp\left(\sum_{k=1}^{\infty}\frac{U_k}{p_i^k}\right)\right]_+.$$

[23] In the computation below, we will use relation (5.A.2) and put $g_{k+1} = 1$.

Let us define

$$\mathcal{G}_r(p_i) \equiv \left[\Lambda^{2N_c-N_f} \frac{\widehat{B}(p_i)}{p_i^{N_c+r}} \exp\left(\sum_{k=1}^{\infty} \frac{U_k}{p_i^k} \right) \right]_+ .$$

By using these relations, let us compute $W'_{\rm cl}$

$$\begin{aligned}
W'_{\rm cl} &= \sum_{r=1}^{N_c} g_r x^{r-1} \\
&= \sum_{r=-\infty}^{N_c} \sum_{i=1}^{l} \sum_{j=0}^{N_c} x^{r-1} p_i^{N_c-j} s_{j-r} L_i \\
&\qquad - \frac{1}{x} \sum_{i=1}^{l} L_i \det(p_i - \Phi) - \sum_{r=1}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} \\
&= \sum_{i=1}^{l} \frac{\det(x-\Phi)}{x-p_i} L_i - \frac{1}{x} \sum_{i=1}^{l} L_i \det(p_i - \Phi) - \sum_{r=1}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} \\
&= \sum_{i=1}^{l} \frac{P_{N_c}(x)}{x-p_i} L_i - \sum_{i=1}^{l} \frac{K(x)}{x-p_i} L_i - \frac{1}{x} \sum_{i=1}^{l} L_i P_{N_c}(p_i) + \frac{1}{x} \sum_{i=1}^{l} L_i K(p_i) \\
&- \sum_{r=1}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} \qquad (5.B.5)
\end{aligned}$$

where we dropped $\mathcal{O}(x^{-2})$. The fifth term above can be written as

$$-\sum_{r=1}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} = -\sum_{r=-\infty}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} + \sum_{i=1}^{l} \frac{1}{x} L_i K(p_i) + \mathcal{O}(x^{-2}).$$

After some manipulation with the factorization form (5.B.3), we obtain a relation

$$\begin{aligned}
Q_{k-n}(x)\sqrt{F_{2n}(x)} &= \frac{P_{N_c}(x)}{\widetilde{H}_{N_c-k}} - \frac{2\Lambda^{2N_c-N_f} \widehat{B}(x)}{\widetilde{H}_{N_c-k} x^N} \exp\left(\sum_{i=1}^{\infty} \frac{U_i}{x^i} \right) \\
&= \frac{P_{N_c}(x)}{\widetilde{H}_{N_c-k}} - \frac{2K(x)}{\widetilde{H}_{N_c-k}} + \mathcal{O}(x^{-2}).
\end{aligned}$$

Substituting this relation into (5.B.5) we obtain

$$\begin{aligned}
W'_{\rm cl} &= Q_{k-n}(x)\sqrt{F_{2n}(x)} - \sum_{i=1}^{l} \frac{1}{x} \left[L_i P_{N_c}(p_i) - 2 L_i K(p_i) \right] + \sum_{i=1}^{l} \frac{K(x)}{x-p_i} L_i \\
&\quad - \sum_{r=-\infty}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} + \mathcal{O}(x^{-2}).
\end{aligned}$$

Finally let us use the following relation [24] ;

$$\sum_{i=1}^{l} \frac{K(x)}{x-p_i} L_i - \sum_{r=-\infty}^{N_c} \sum_{i=1}^{l} L_i \mathcal{G}_r(p_i) x^{r-1} = \mathcal{O}(x^{-2}). \qquad (5.B.6)$$

After all, by squaring $W'_{\rm cl}$, we have

$$Q_{k-n}(x)^2 F_{2n}(x) = W'^2_{\rm cl}(x) + 2\sum_{i=1}^{l} \frac{1}{x}\left[L_i P_{N_c}(p_i) - 2L_i K(p_i)\right] x^{k-1} + \mathcal{O}(x^{k-2}),$$

$$b_{k-1} = 2\sum_{i=1}^{l} \frac{1}{x}\left[L_i P_{N_c}(p_i) - 2L_i K(p_i)\right].$$

On the other hand the variation of $W_{\rm low}$ with respect to Λ is given by

$$\frac{\partial W_{\rm low}}{\partial \log \Lambda^{2N_c-N_f}} = \sum_{i=1}^{l} L_i K(p_i) - \frac{1}{2}\sum_{i=1}^{l} L_i P_{N_c}(p_i) = -\frac{b_{k-1}}{4}. \qquad (5.B.7)$$

This is one of the main results for our proof. In the dual geometry analysis below, we will see the similar relation.

In the classical limit, we have only to consider the expectation value of Φ. In our assumption, gauge symmetry breaks as $U(N_c) \to \prod_i^n U(N_{c,i})$ we have

[24]For simplicity, we ignore $\sum_{i=1}^{l} L_i$. To prove the relation, let us consider the circle integral over C. Until now, since we assumed $p_i < x$, the point p_i should be included in the contour C. Multiplying $\frac{1}{x^k}$ $k \geq 0$ and taking the circle integral, we can pick up the coefficient c_{k-1} of x^{k-1}. In addition, if we denote a polynomial $M \equiv \Lambda^{2N_c-N_f} \frac{\widehat{B}(x)}{x^{N_c}} \exp\left(\sum_{i=1}^{\infty} \frac{U_i}{x^i}\right) \equiv \sum_{j=-\infty}^{\infty} a_j x^j$ we can obtain following relation,

$$\frac{[M]_+}{x^k} = \left[\frac{M}{x^k}\right]_+ + \sum_{j=0}^{k-1} a_j x^{-(k-j)} \iff \frac{K(x)}{x^k} = \mathcal{G}_k(x) + \sum_{j=0}^{k-1} a_j x^{-(k-j)}$$

where right hand side means circle integral of left hand side. Thus, we obtain the c_{k-1} as

$$\begin{aligned} c_{k-1} &= \sum_{j=0}^{k-1} \oint_{x=0} \frac{a_j x^{j-k}}{x-p_i} dx + \sum_{j=0}^{k-1} \oint_{x=p_i} \frac{a_j x^{j-k}}{x-p_i} dx \\ &= -\sum_{j=0}^{k-1} \oint_{x=0} a_j x^{j-k} \sum_{n=1}^{\infty} \frac{x^n}{p_i^{n+1}} dx + \sum_{j=0}^{k-1} a_j p_i^{j-k} = 0 \end{aligned}$$

where we used that around $x=0$, $\frac{1}{x-p_i} = -\sum_{n=1}^{\infty} \frac{x^n}{p_i^{n+1}}$. Therefore the left hand side of (5.B.6) can be written as $\sum_{j=-\infty}^{\infty} c_j x^j = \sum_{j=-\infty}^{-2} c_j x^j = \mathcal{O}(x^{-2})$.

$\mathrm{Tr}\,\Phi = \sum_{i=1}^{n} N_{c,i} a_i$. Therefore in the classical limit W_{low} behaves as

$$W_{\mathrm{cl}} = \sum_{i}^{n} N_{c,i} W(a_i). \tag{5.B.8}$$

By comparing (5.B.7) and (5.B.8) and the similar result which we will see in the dual geometry analysis below, we will show (5.B.1) and (5.B.2). Let us move to the dual geometry analysis.

5.B.2 Dual geometry analysis with some closed branch cuts

As we have already seen in the main text, a solution with $\langle S_i \rangle = 0$ appears for some special choice of z_I, the position of flavor poles. In our present proof, however, we put some of S_i to be zero from the beginning, *without specifying* z_I. More precisely, what we prove in this Appendix is as follows: for a given choice of z_I, if there exists a solution to the factorization problem with some of $\langle S_i \rangle$ vanishing, then we can construct a dual geometry which gives the same low energy effective superpotential as the one given by the solution to the factorization problem, by setting some of S_i to zero from the beginning. Therefore, this analysis *does not* tell us when a solution with $\langle S_i \rangle = 0$ appears. To know that within the matrix model formalism, we have to go back to string theory and consider an explanation such as the one given in [87].

Now let us start our proof. Again, let k be the degree of tree level superpotential $W'_{\mathrm{tree}}(x)$ and n be the number of $U(1)$ at low energy. To realize this situation, we need to consider that $(k-n)$ branch cuts on the Riemann surface should be closed, which corresponds to $\langle S_i \rangle = 0$. Here, there is one important thing: As we know from the expansion of W_{eff} in terms of Λ (e.g. see (5.4.13)), we cannot obtain any solutions with $\langle S_i \rangle = 0$ if we assume that S_i is dynamical and solve its equation of motion. Therefore to realize the situation with vanishing $\langle S_i \rangle$, we

must put $S_i = 0$ at the off-shell level by hand. With this in mind, let us study dual geometry which corresponds to the gauge theories above. In the field theory, we assumed that the Riemann surface had $(k-n)$ closed branch cuts. Thus, in this dual geometry analysis we must assume that at off-shell level, $(k-n)$ S_i's must be zero. For convenience, we assume that first n S_i's are non-zero and the remaining $(k-n)$ vanish,

$$S_i \neq 0, \quad i = 1, \cdots n, \quad S_i = 0, \quad i = n+1, \cdots k.$$

Therefore the Riemann surface can be written as

$$y^2 = F_{2n}(x) Q_{k-n}^2 = W'(x)^2 + b_{k-1} x^{k-1} + \cdots$$

The effective superpotential in dual geometry corresponding to $U(N_c)$ gauge theory with N_f flavors was given in [35] (See also (5.2.12)) and in the classical limit it behaves as [25]

$$W_{\text{eff}}\Big|_{\text{cl}} = \sum_{i=1}^{n} N_{c,i} W(a_i). \tag{5.B.9}$$

As discussed in the previous section, existence of flavors does not change the Riemann surface $y(x)$. In other words, Riemann surface is not singular at x_I (roots of $B(x)$),

$$\oint_{x_I} y(x) dx = 0, \qquad y(x) = \sqrt{W'(x)^2 + b_{k-1} x^{k-1} + \cdots}.$$

Therefore as in [76], by deforming contours of all S_i's and evaluating the residue at infinity on the first sheet, we obtain the following relation,

$$\sum_{i=1}^{n} S_i = \sum_{i=1}^{k} S_i = -\frac{1}{4} b_{k-1},$$

[25] Remember that in this Appendix we are assuming only Coulomb branch. For the Higgs branch, see (7.11) and (7.12) in [35].

where we used $\sum_{i=n+1}^{k} S_i = 0$.

With this relation in mind, next we consider the variation of W_{eff} with respect to S_i:

$$\frac{\partial W_{\text{eff}}(S_i, \Lambda)}{\partial S_i} = 0, \quad i = 1 \cdots n. \tag{5.B.10}$$

Solving these equations, we obtain the expectation values, $\langle S_i \rangle$. Of course, these vacuum expectation values depend on Λ, g_r and z_I. Thus when we evaluate the variation of $W_{\text{eff}}(\langle S_i \rangle, \Lambda)$ with respect to Λ, we have to pay attention to implicit dependence on Λ. However the implicit dependence does not contribute because of the equation of motion (5.B.10):

$$\begin{aligned} \frac{dW_{\text{eff}}(\langle S_i \rangle, \Lambda)}{d\Lambda} &= \sum_{i=1}^{n} \frac{\partial \langle S_i \rangle}{\partial \Lambda} \cdot \frac{\partial W_{\text{eff}}(\langle S_i \rangle, \Lambda)}{\partial \langle S_i \rangle} + \frac{\partial W_{\text{eff}}(\langle S_i \rangle, \Lambda)}{\partial \Lambda} \\ &= \frac{\partial W_{\text{eff}}(\langle S_i \rangle, \Lambda)}{\partial \Lambda}. \end{aligned}$$

On the other hand, explicit dependence on Λ can be easily obtained by monodromy analysis. Here let us recall the fact that the presence of fundamentals does not change the Riemann surface. In fact, looking at (5.2.12) we can read off the dependence from the term $2\pi i \tau_0 = \log \left(\frac{B_L \Lambda^{2N_c - N_f}}{\Lambda_0^{2N_c - L}} \right)$,

$$\frac{dW_{\text{eff}}(\langle S_i \rangle, \Lambda)}{d \log \Lambda^{2N_c - N_f}} = S = -\frac{b_{k-1}}{4}. \tag{5.B.11}$$

To finish our proof, we have to pay attention to $f_{k-1}(x)$, on-shell. Namely putting $\langle S_i \rangle$ into $f_{k-1}(x)$ what kind of property does it have? To see it, let us consider change of variables from S_i's to b_i's. As discussed in [17] the Jacobian of the change is non-singular if $0 \leq j \leq k-2$,

$$\frac{\partial S_i}{\partial b_j} = -\frac{1}{8\pi i} \oint_{A_i} dx \frac{x^j}{\sqrt{W'(x)^2 + f(x)}}.$$

In our present case, since only n of k S_i's are dynamical variable, we use b_i, $i = 0, \cdots n-1$ in a function $f_{k-1}(x) = \sum b_i x^i$ as new variables, instead of S_i's.

As discussed in [76, 80, 35], by using Abel's theorem, the equation of motions for b_i's is interpreted as an existence condition of a meromorphic function that has an N_c-th order pole at infinity on the first sheet and an $(N_c - N_f)$-th order zero at infinity on the second sheet of Σ and a first order zero at $\widetilde{q}_I$. For a theory with $N_f \leq 2N_c$, such a function can be constructed as follows [53, 35]:

$$\psi(x) = P_{N_c}(x) + \sqrt{P_{N_c}^2(x) - 4\Lambda^{2N_c - N_f}\widehat{B}(x)}.$$

For this function to be single valued on the matrix model curve $y(x)$, the following condition must be satisfied,

$$\begin{aligned} P_{N_c}^2(x) - 4\Lambda^{2N_c-N_f}\widehat{B}(x) &= F_{2n}(x)H_{N_c-n}^2(x) \\ W'(x)^2 + f(x) &= F_{2n}(x)Q_{k-n}^2(x) \end{aligned}$$

This is exactly the same as the factorization form we already see in the field theory analysis. Therefore the value b_{k-1} of on-shell matrix model curve in dual theory is the same one for field theory analysis. Comparing two results, (5.B.7) and (5.B.8) with corresponding results for the dual geometry analysis, (5.B.9) and (5.B.11) we have shown the equivalence between these two descriptions of effective superpotentials.

5.C Computation of superpotential — gauge theory side

In this Appendix, we demonstrate the factorization method used in subsection 5.4.1 to compute the low energy superpotential, taking the $N_f = 4$ case as an example. Therefore there are two kinds of solutions for the factorization problem (5.4.5) and (5.4.6).

• **The breaking pattern** $\widehat{U(2)} \times U(1)$

The first kind of solution for the factorization problem is given by

$$A = 0, \qquad B = 2a, \qquad C = a^2 - 4\Lambda^2, \qquad D = 0, \qquad c = 0, \qquad b = 0.$$

In the classical limit $\Lambda \to 0$, we can see the characteristic function goes as $P_3(x) \to x^2\,(x+a)$, which means that the breaking pattern is $\widehat{U(2)} \times U(1)$. Note that since we are assuming $m_f = 0$, the notation " $\widehat{\ }$ " should be used for the gauge group that corresponds to the cut near the critical point at $x = 0$. Inserting these solutions into (5.4.7) we obtain one constraint,

$$\Delta^2 = a^2 + 8\Lambda^2.$$

We can easily represent a as a Taylor expansion of Λ:

$$a = -1 + 4\,T + 8\,T^2 + 32\,T^3 + 160\,T^4 + 896\,T^5 + 5376\,T^6 + \cdots,$$

where we put $\Delta = 1$ and defined $T \equiv \Lambda^2$. [26] The coefficients of $P_3(x)$ are related to the Casimirs $U_j = \frac{1}{j}\langle \mathrm{Tr}[\Phi^j] \rangle$ as follows. For $N_c = 3$, $N_f = 4$, (5.A.1) reads

$$\begin{aligned} P_3(x) &= x^3 \exp\left(-\sum_{j=1}^{\infty} \frac{U_j}{x^j} \right) + \Lambda^5 \frac{x^4}{x^3} \exp\left(\sum_{j=1}^{\infty} \frac{U_j}{x^j} \right) \\ &= x^3 - U_1 x^2 + \left(-U_2 + \frac{U_1^2}{2} + \Lambda^2 \right) x \\ &\qquad + \left(-U_3 + U_1 U_2 - \frac{U_1^3}{6} - \Lambda^2 U_1 \right) + \cdots . \end{aligned}$$

Comparing the coefficients, we obtain

$$U_1 = -a, \qquad U_2 = -b + \frac{a^2}{2} + \Lambda^2, \qquad U_3 = -c + ab - \frac{a^3}{3} - a\Lambda^2.$$

[26] If we take care of a phase factor of Λ, we will obtain the effective superpotentials corresponding to each vacuum. However in our present calculation, we want to check whether the effective superpotentials of two method, field theory and dual geometry, agree with each other. Therefore, we have only to pay attention to the coefficients in W_{low}, neglecting the phase factor.

Furthermore, one can compute $a_{1,2}$ from (5.4.7). Plugging all these into (5.4.9), we finally obtain

$$W_{\text{low}} = W_{\text{cl}} + T - 6T^2 - \frac{40T^3}{3} - 56T^4 - 288T^5 - \frac{4928T^6}{3} + \cdots,$$
$$T \equiv \Lambda^2, \qquad W_{\text{cl}} = -\frac{1}{6}.$$

• **The breaking pattern** $\widehat{U(1)} \times U(2)$

The other kind of solution for the factorization problem is given by

$$A = \frac{1}{2}(-a - 2\eta\Lambda), \qquad B = a - 2\eta\Lambda, \qquad C = \frac{1}{4}(a + 2\eta\Lambda)^2,$$
$$D = 0, \qquad c = 0, \qquad b = \frac{1}{4}(a + 2\eta\Lambda)^2$$

where $\eta \equiv \pm 1$. These solutions correspond to the breaking pattern $\widehat{U(1)} \times U(2)$ in the classical limit. Inserting these solutions into (5.4.7) we obtain one constraint,

$$\Delta^2 = \frac{1}{4}(a^2 - 20a\,\eta\,\Lambda + 4\Lambda^2).$$

Again, let us represent a as a Taylor series of Λ:

$$a = -2 + 10\,T - 24\,T^2 + 144\,T^4 - 1728\,T^6 + \cdots,$$

where we put $\Delta = 1$, $\eta = 1$ and defined $T \equiv \Lambda$. Doing the same way as previous breaking pattern, we can compute the effective superpotential as

$$W_{\text{low}} = W_{\text{cl}} + 2T - 13T^2 + \frac{176T^3}{3} - 138T^4 + 792T^6 + \cdots,$$
$$T \equiv \Lambda, \qquad W_{\text{cl}} = -\frac{1}{3}.$$

The other cases with $N_f = 1, 2, 3$ and 5 can be done analogously.

References

[1] S. Weinberg, "The Quantum Theory Of Fields. Vol. 3: Supersymmetry,"

[2] M. E. Peskin, "Duality in supersymmetric Yang-Mills theory," arXiv:hep-th/9702094.

[3] P. Argyres, "Introduction to Global Supersymmetry" lecture notes, `http://www.physics.uc.edu/~argyres/661/index.html`

[4] K. A. Intriligator and N. Seiberg, "Lectures on supersymmetric gauge theories and electric-magnetic duality," Nucl. Phys. Proc. Suppl. **45BC**, 1 (1996) [arXiv:hep-th/9509066].

[5] R. Dijkgraaf and C. Vafa, "Matrix models, topological strings, and supersymmetric gauge theories," Nucl. Phys. B **644**, 3 (2002) [arXiv:hep-th/0206255].

[6] R. Dijkgraaf and C. Vafa, "On geometry and matrix models," Nucl. Phys. B **644**, 21 (2002) [arXiv:hep-th/0207106].

[7] R. Dijkgraaf and C. Vafa, "A perturbative window into non-perturbative physics," arXiv:hep-th/0208048.

[8] R. Gopakumar, "N = 1 theories and a geometric master field," JHEP **0305**, 033 (2003) [arXiv:hep-th/0211100].

[9] C. Vafa, "Superstrings and topological strings at large N," J. Math. Phys. **42**, 2798 (2001) [arXiv:hep-th/0008142].

[10] F. Cachazo, K. A. Intriligator and C. Vafa, "A large N duality via a geometric transition," Nucl. Phys. B **603**, 3 (2001) [arXiv:hep-th/0103067].

[11] M. Bershadsky, S. Cecotti, H. Ooguri and C. Vafa, "Kodaira-Spencer theory of gravity and exact results for quantum string amplitudes," Commun. Math. Phys. **165**, 311 (1994) [arXiv:hep-th/9309140].

[12] G. 't Hooft, "A Two-Dimensional Model For Mesons," Nucl. Phys. B **75**, 461 (1974).

[13] R. Dijkgraaf, M. T. Grisaru, C. S. Lam, C. Vafa and D. Zanon, "Perturbative computation of glueball superpotentials," Phys. Lett. B **573**, 138 (2003) [arXiv:hep-th/0211017].

[14] H. Ita, H. Nieder and Y. Oz, "Perturbative computation of glueball superpotentials for SO(N) and USp(N)," JHEP **0301**, 018 (2003) [arXiv:hep-th/0211261].

[15] M. Aganagic, K. Intriligator, C. Vafa and N. P. Warner, "The glueball superpotential," arXiv:hep-th/0304271.

[16] J. Wess and J. Bagger, "Supersymmetry And Supergravity,"

[17] F. Cachazo, M. R. Douglas, N. Seiberg and E. Witten, "Chiral rings and anomalies in supersymmetric gauge theory," JHEP **0212**, 071 (2002) [arXiv:hep-th/0211170].

[18] N. Seiberg, "Adding fundamental matter to 'Chiral rings and anomalies in supersymmetric gauge theory'," JHEP **0301**, 061 (2003) [arXiv:hep-th/0212225].

[19] R. Argurio, G. Ferretti and R. Heise, "An introduction to supersymmetric gauge theories and matrix models," arXiv:hep-th/0311066.

[20] K. Konishi, "Anomalous Supersymmetry Transformation Of Some Composite Operators In Sqcd," Phys. Lett. B **135**, 439 (1984); K. i. Konishi and K. i. Shizuya, "Functional Integral Approach To Chiral Anomalies In Supersymmetric Gauge Theories," Nuovo Cim. A **90**, 111 (1985).

[21] P. L. Cho and P. Kraus, "Symplectic SUSY gauge theories with antisymmetric matter," Phys. Rev. D **54**, 7640 (1996) [arXiv:hep-th/9607200].

[22] C. Csaki, W. Skiba and M. Schmaltz, "Exact results and duality for Sp(2N) SUSY gauge theories with an antisymmetric tensor," Nucl. Phys. B **487**, 128 (1997) [arXiv:hep-th/9607210].

[23] P. Kraus and M. Shigemori, "On the matter of the Dijkgraaf-Vafa conjecture," JHEP **0304**, 052 (2003) [arXiv:hep-th/0303104].

[24] F. Cachazo, "Notes on supersymmetric Sp(N) theories with an antisymmetric tensor," arXiv:hep-th/0307063.

[25] C. h. Ahn, B. Feng and Y. Ookouchi, "Phases of $N = 1$ USp(2N(c)) gauge theories with flavors," Phys. Rev. D **69**, 026006 (2004) [arXiv:hep-th/0307190].

[26] M. Matone, "The affine connection of supersymmetric SO(N)/Sp(N) theories," JHEP **0310**, 068 (2003) [arXiv:hep-th/0307285].

[27] S. G. Naculich, H. J. Schnitzer and N. Wyllard, "Matrix-model description of N = 2 gauge theories with non-hyperelliptic Seiberg-Witten curves," Nucl. Phys. B **674**, 37 (2003) [arXiv:hep-th/0305263].

[28] M. Gomez-Reino, S. G. Naculich and H. J. Schnitzer, "Improved matrix-model calculation of the N = 2 prepotential," JHEP **0404**, 033 (2004) [arXiv:hep-th/0403129].

[29] K. Landsteiner and C. I. Lazaroiu, "On Sp(0) factors and orientifolds," Phys. Lett. B **588**, 210 (2004) [arXiv:hep-th/0310111].

[30] L. F. Alday and M. Cirafici, "Effective superpotentials via Konishi anomaly," JHEP **0305**, 041 (2003) [arXiv:hep-th/0304119].

[31] P. Kraus, A. V. Ryzhov and M. Shigemori, "Loop equations, matrix models, and N = 1 supersymmetric gauge theories," JHEP **0305**, 059 (2003) [arXiv:hep-th/0304138].

[32] J. M. Maldacena, "The large N limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231 (1998) [Int. J. Theor. Phys. **38**, 1113 (1999)] [arXiv:hep-th/9711200].

[33] E. Witten, "Anti-de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253 (1998) [arXiv:hep-th/9802150].

[34] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, "Gauge theory correlators from non-critical string theory," Phys. Lett. B **428**, 105 (1998) [arXiv:hep-th/9802109].

[35] F. Cachazo, N. Seiberg and E. Witten, "Chiral Rings and Phases of Supersymmetric Gauge Theories," JHEP **0304**, 018 (2003) [arXiv:hep-th/0303207].

[36] N. Seiberg and E. Witten, "Electric - magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory," Nucl. Phys. B **426**, 19 (1994) [Erratum-ibid. B **430**, 485 (1994)] [arXiv:hep-th/9407087].

[37] N. Seiberg and E. Witten, "Monopoles, duality and chiral symmetry breaking in N=2 supersymmetric QCD," Nucl. Phys. B **431**, 484 (1994) [arXiv:hep-th/9408099].

[38] G. Veneziano and S. Yankielowicz, "An Effective Lagrangian For The Pure N=1 Supersymmetric Yang-Mills Theory," Phys. Lett. B **113**, 231 (1982).

[39] S. K. Ashok, R. Corrado, N. Halmagyi, K. D. Kennaway and C. Romelsberger, "Unoriented strings, loop equations, and N = 1 superpotentials from matrix models," Phys. Rev. D **67**, 086004 (2003) [arXiv:hep-th/0211291].

[40] R. A. Janik and N. A. Obers, "SO(N) superpotential, Seiberg-Witten curves and loop equations," Phys. Lett. B **553**, 309 (2003) [arXiv:hep-th/0212069].

[41] V. Balasubramanian, J. de Boer, B. Feng, Y. H. He, M. x. Huang, V. Jejjala and A. Naqvi, "Multi-trace superpotentials vs. matrix models," Commun. Math. Phys. **242**, 361 (2003) [arXiv:hep-th/0212082].

[42] R. Argurio, V. L. Campos, G. Ferretti and R. Heise, "Baryonic corrections to superpotentials from perturbation theory," Phys. Lett. B **553**, 332 (2003) [arXiv:hep-th/0211249];

[43] K. A. Intriligator, "'Integrating in' and exact superpotentials in 4-d," Phys. Lett. B **336**, 409 (1994) [arXiv:hep-th/9407106].

[44] E. Witten, "Chiral Ring Of Sp(N) And SO(N) Supersymmetric Gauge Theory In Four Dimensions," arXiv:hep-th/0302194.

[45] E. Brezin, C. Itzykson, G. Parisi and J. B. Zuber, "Planar Diagrams," Commun. Math. Phys. **59**, 35 (1978).

[46] S. G. Naculich, H. J. Schnitzer and N. Wyllard, "Cubic curves from matrix models and generalized Konishi anomalies," JHEP **0308**, 021 (2003) [arXiv:hep-th/0303268].

[47] N. Seiberg, "Adding fundamental matter to 'Chiral rings and anomalies in supersymmetric gauge theory'," JHEP **0301**, 061 (2003) [arXiv:hep-th/0212225].

[48] R. Casero and E. Trincherini, "Quivers via anomaly chains," JHEP **0309**, 041 (2003) [arXiv:hep-th/0304123].

[49] A. Brandhuber, H. Ita, H. Nieder, Y. Oz and C. Romelsberger, "Chiral rings, superpotentials and the vacuum structure of N = 1 supersymmetric gauge theories," Adv. Theor. Math. Phys. **7**, 269 (2003) [arXiv:hep-th/0303001].

[50] A. Klemm, K. Landsteiner, C. I. Lazaroiu and I. Runkel, "Constructing gauge theory geometries from matrix models," JHEP **0305**, 066 (2003) [arXiv:hep-th/0303032].

[51] E. Witten, "An SU(2) Anomaly," Phys. Lett. B **117**, 324 (1982).

[52] S. G. Naculich, H. J. Schnitzer and N. Wyllard, "The N = 2 U(N) gauge theory prepotential and periods from a perturbative matrix model calculation," Nucl. Phys. B **651**, 106 (2003) [arXiv:hep-th/0211123].

[53] S. G. Naculich, H. J. Schnitzer and N. Wyllard, "Matrix model approach to the N = 2 U(N) gauge theory with matter in the fundamental representation," JHEP **0301**, 015 (2003) [arXiv:hep-th/0211254].

[54] N. Dorey, T. J. Hollowood and S. P. Kumar, "S-duality of the Leigh-Strassler deformation via matrix models," JHEP **0212**, 003 (2002) [arXiv:hep-th/0210239].

[55] F. Ferrari, "On exact superpotentials in confining vacua," Nucl. Phys. B **648**, 161 (2003) [arXiv:hep-th/0210135].

[56] R. Gopakumar and C. Vafa, "On the gauge theory/geometry correspondence," Adv. Theor. Math. Phys. **3**, 1415 (1999) [arXiv:hep-th/9811131].

[57] R. Dijkgraaf and C. Vafa, "N = 1 supersymmetry, deconstruction, and bosonic gauge theories," arXiv:hep-th/0302011.

[58] N. Dorey, T. J. Hollowood, S. Prem Kumar and A. Sinkovics, "Exact superpotentials from matrix models," JHEP **0211**, 039 (2002) [arXiv:hep-th/0209089].

[59] F. Cachazo, S. Katz and C. Vafa, "Geometric transitions and N = 1 quiver theories," arXiv:hep-th/0108120.

[60] A. Strominger, "Massless black holes and conifolds in string theory," Nucl. Phys. B **451**, 96 (1995) [arXiv:hep-th/9504090].

[61] T. R. Taylor and C. Vafa, "RR flux on Calabi-Yau and partial supersymmetry breaking," Phys. Lett. B **474**, 130 (2000) [arXiv:hep-th/9912152].

[62] P. Mayr, "On supersymmetry breaking in string theory and its realization in brane worlds," Nucl. Phys. B **593**, 99 (2001) [arXiv:hep-th/0003198].

[63] S. Sinha and C. Vafa, "SO and Sp Chern-Simons at large N," arXiv:hep-th/0012136.

[64] J. D. Edelstein, K. Oh and R. Tatar, "Orientifold, geometric transition and large N duality for SO/Sp gauge theories," JHEP **0105**, 009 (2001) [arXiv:hep-th/0104037].

[65] K. Landsteiner, C. I. Lazaroiu and R. Tatar, "(Anti)symmetric matter and superpotentials from IIB orientifolds," JHEP **0311**, 044 (2003) [arXiv:hep-th/0306236].

[66] C. Csaki, M. Schmaltz, W. Skiba and J. Terning, "Gauge theories with tensors from branes and orientifolds," Phys. Rev. D **57**, 7546 (1998) [arXiv:hep-th/9801207].

[67] K. Landsteiner, C. I. Lazaroiu and R. Tatar, "Chiral field theories from conifolds," JHEP **0311**, 057 (2003) [arXiv:hep-th/0310052].

[68] E. Witten, "Baryons and branes in anti de Sitter space," JHEP **9807**, 006 (1998) [arXiv:hep-th/9805112].

[69] D. J. Gross and H. Ooguri, "Aspects of large N gauge theory dynamics as seen by string theory," Phys. Rev. D **58**, 106002 (1998) [arXiv:hep-th/9805129].

[70] S. S. Gubser and I. R. Klebanov, "Baryons and domain walls in an N = 1 superconformal gauge theory," Phys. Rev. D **58**, 125025 (1998) [arXiv:hep-th/9808075].

[71] H. Fuji and Y. Ookouchi, "Confining phase superpotentials for SO/Sp gauge theories via geometric transition," JHEP **0302**, 028 (2003) [arXiv:hep-th/0205301].

[72] R. Abbaspur, A. Imaanpur and S. Parvizi, "N = 2 SO(N) SYM theory from matrix model," JHEP **0307**, 043 (2003) [arXiv:hep-th/0302083].

[73] K. A. Intriligator, "New RG fixed points and duality in supersymmetric SP(N(c)) and SO(N(c)) gauge theories," Nucl. Phys. B **448**, 187 (1995) [arXiv:hep-th/9505051].

[74] R. Dijkgraaf, S. Gukov, V. A. Kazakov and C. Vafa, "Perturbative analysis of gauged matrix models," Phys. Rev. D **68**, 045007 (2003) [arXiv:hep-th/0210238].

[75] C. Csaki and H. Murayama, "Instantons in partially broken gauge groups," Nucl. Phys. B **532**, 498 (1998) [arXiv:hep-th/9804061].

[76] F. Cachazo and C. Vafa, "N = 1 and N = 2 geometry from fluxes," arXiv:hep-th/0206017.

[77] S. Elitzur, A. Forge, A. Giveon, K. A. Intriligator and E. Rabinovici, "Massless Monopoles Via Confining Phase Superpotentials," Phys. Lett. B **379**, 121 (1996) [arXiv:hep-th/9603051].

[78] M. L. Mehta, "Random matrices," Academic Press (1991).

[79] F. Cachazo, N. Seiberg and E. Witten, "Phases of N = 1 supersymmetric gauge theories and matrices," JHEP **0302**, 042 (2003) [arXiv:hep-th/0301006].

[80] Y. Ookouchi, "N = 1 gauge theory with flavor from fluxes," JHEP **0401**, 014 (2004) [arXiv:hep-th/0211287].

[81] V. Balasubramanian, B. Feng, M. x. Huang and A. Naqvi, "Phases of N = 1 supersymmetric gauge theories with flavors," Annals Phys. **310**, 375 (2004) [arXiv:hep-th/0303065].

[82] B. Feng, "Geometric dual and matrix theory for SO/Sp gauge theories," Nucl. Phys. B **661**, 113 (2003) [arXiv:hep-th/0212010].

[83] C. h. Ahn and Y. Ookouchi, "Phases of N = 1 supersymmetric SO / Sp gauge theories via matrix model," JHEP **0303**, 010 (2003) [arXiv:hep-th/0302150].

[84] C. h. Ahn, B. Feng and Y. Ookouchi, "Phases of N = 1 SO(N(c)) gauge theories with flavors," Nucl. Phys. B **675**, 3 (2003) [arXiv:hep-th/0306068].

[85] K. A. Intriligator and N. Seiberg, "Duality, monopoles, dyons, confinement and oblique confinement in supersymmetric SO(N(c)) gauge theories," Nucl. Phys. B **444**, 125 (1995) [arXiv:hep-th/9503179].

[86] M. R. Douglas and S. H. Shenker, "Dynamics of SU(N) supersymmetric gauge theory," Nucl. Phys. B **447**, 271 (1995) [arXiv:hep-th/9503163].

[87] K. Intriligator, P. Kraus, A. V. Ryzhov, M. Shigemori and C. Vafa, "On low rank classical groups in string theory, gauge theory and matrix models," Nucl. Phys. B **682**, 45 (2004) [arXiv:hep-th/0311181].

[88] P. Svrcek, "Chiral rings, vacua and gaugino condensation of supersymmetric gauge theories," arXiv:hep-th/0308037.

[89] P. C. Argyres, M. R. Plesser and N. Seiberg, "The Moduli Space of N=2 SUSY QCD and Duality in N=1 SUSY QCD," Nucl. Phys. B **471**, 159 (1996) [arXiv:hep-th/9603042].

[90] G. Carlino, K. Konishi and H. Murayama, "Dynamical symmetry breaking in supersymmetric SU(n(c)) and USp(2n(c)) gauge theories," Nucl. Phys. B **590**, 37 (2000) [arXiv:hep-th/0005076].

[91] G. Carlino, K. Konishi, S. P. Kumar and H. Murayama, "Vacuum structure and flavor symmetry breaking in supersymmetric SO(n(c)) gauge theories," Nucl. Phys. B **608**, 51 (2001) [arXiv:hep-th/0104064].

[92] Y. Demasure and R. A. Janik, "Explicit factorization of Seiberg-Witten curves with matter from random matrix models," Nucl. Phys. B **661**, 153 (2003) [arXiv:hep-th/0212212].

[93] K. A. Intriligator and N. Seiberg, "Phases of N=1 supersymmetric gauge theories in four-dimensions," Nucl. Phys. B **431**, 551 (1994) [arXiv:hep-th/9408155].

[94] Y. Demasure, "Affleck-Dine-Seiberg from Seiberg-Witten," [arXiv:hep-th/0307082].

[95] J. Erlich, S. Hong and M. Unsal, "Matrix models, monopoles and modified moduli," [arXiv:hep-th/0312054].

[96] N. Seiberg, "Electric - magnetic duality in supersymmetric nonAbelian gauge theories," Nucl. Phys. B **435**, 129 (1995) [arXiv:hep-th/9411149].

[97] C. h. Ahn and Y. Ookouchi, "The matrix model curve near the singularities," Phys. Rev. D **69**, 086006 (2004) [arXiv:hep-th/0309156].

[98] C. h. Ahn, B. Feng, Y. Ookouchi and M. Shigemori, "Supersymmetric gauge theories with flavors and matrix models," Nucl. Phys. B **698**, 3 (2004) [arXiv:hep-th/0405101].

[99] I. Bena, R. Roiban and R. Tatar, "Baryons, boundaries and matrix models," Nucl. Phys. B **679**, 168 (2004) [arXiv:hep-th/0211271].

[100] I. Bena, H. Murayama, R. Roiban and R. Tatar, "Matrix model description of baryonic deformations," JHEP **0305**, 049 (2003) [arXiv:hep-th/0303115].